

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:09 ; Search time 19.0439 Seconds

(without alignments)
1661.397 Million cell updates/sec

Title: US-10-085-167-2

Perfect score: 1726

Sequence: 1 MLPLLLGLGPAACWALGFT.....LVYDDLAPAAPGLGASELL 329

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR 76:*
1: pirl:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB ID | Description |
|------------|-------|---------------|--------|----------|--------------------|
| 1 | 213 | 12.3 | 219 | 2 T14782 | hypothetical prote |
| 2 | 196 | 11.4 | 244 | 2 JC4708 | gelatin-binding 28 |
| 3 | 184 | 10.7 | 215 | 2 B48150 | hibernation-relate |
| 4 | 182 | 10.5 | 253 | 2 C1HUQB | complement subcomp |
| 5 | 176.5 | 10.2 | 253 | 2 S49158 | complement protein |
| 6 | 175 | 10.1 | 744 | 1 S23298 | collagen alpha 1(V |
| 7 | 173 | 10.0 | 744 | 1 A34246 | collagen alpha 1(V |
| 8 | 172 | 10.0 | 245 | 1 C1HUQC | complement subcomp |
| 9 | 170 | 9.8 | 744 | 2 S15435 | collagen alpha 1(V |
| 10 | 169.5 | 9.8 | 674 | 2 S23297 | collagen alpha 1(X |
| 11 | 168.5 | 9.8 | 680 | 2 S31216 | collagen alpha 1(X |
| 12 | 168 | 9.7 | 245 | 2 S19018 | complement subcomp |
| 13 | 166.5 | 9.6 | 680 | 1 CGHU1D | collagen alpha 1(X |
| 14 | 165.5 | 9.6 | 743 | 1 S23779 | collagen alpha 1(V |
| 15 | 164.5 | 9.5 | 253 | 2 I49560 | complement C1q B c |
| 16 | 163.5 | 9.5 | 674 | 2 S13301 | collagen alpha 1(X |
| 17 | 163 | 9.4 | 224 | 2 A60032 | cerebellin-like g1 |
| 18 | 160.5 | 9.3 | 246 | 2 S29328 | complement subcomp |
| 19 | 158 | 9.2 | 193 | 2 A37873 | cerebellin precurs |
| 20 | 151.5 | 8.8 | 245 | 1 C1HUQA | complement subcomp |
| 21 | 150 | 8.7 | 196 | 2 A48150 | hibernation-relate |
| 22 | 147.5 | 8.5 | 215 | 2 C48150 | multimerin, endoth |
| 23 | 147 | 8.5 | 1228 | 2 A57384 | collagen precursor |
| 24 | 145.5 | 8.4 | 423 | 2 A55797 | collagen alpha 2(V |
| 25 | 139.5 | 8.1 | 635 | 2 A57131 | collagen alpha 2(V |
| 26 | 135.5 | 7.9 | 170 | 2 B57131 | hypothetical prote |
| 27 | 131.5 | 7.6 | 992 | 2 T08772 | Basoon protein - |
| 28 | 98 | 5.7 | 3938 | 2 T42761 | 3-isopropylmalate |
| 29 | 97.5 | 5.6 | 476 | 2 T29083 | |

| | | | | | |
|----|------|-----|-------|----------|--------------------|
| 30 | 95 | 5.5 | 495 | 1 G64878 | aldehyde dehydroge |
| 31 | 94.5 | 5.5 | 619 | 2 A87499 | hypothetical prote |
| 32 | 94 | 5.4 | 495 | 2 E90863 | aldehyde dehydroge |
| 33 | 94 | 5.4 | 495 | 2 D85755 | aldehyde dehydroge |
| 34 | 94 | 5.4 | 810 | 2 B84185 | cytochrome-like pr |
| 35 | 94 | 5.4 | 913 | 2 AB2587 | translation initia |
| 36 | 94 | 5.4 | 913 | 2 B97369 | translation initia |
| 37 | 94 | 5.4 | 1045 | 2 A39199 | endoglucanase B (E |
| 38 | 94 | 5.4 | 7463 | 2 T36248 | CDA peptide synthe |
| 39 | 93.5 | 5.4 | 686 | 1 VGBENA | glycoprotein H pre |
| 40 | 93 | 5.4 | 462 | 2 B75306 | probable cycloprop |
| 41 | 93 | 5.4 | 481 | 2 B75480 | hypothetical prote |
| 42 | 93 | 5.4 | 1239 | 1 QOBE10 | BOLFI protein - hu |
| 43 | 92.5 | 5.4 | 420 | 2 E75130 | hypothetical prote |
| 44 | 92.5 | 5.4 | 1334 | 2 T50568 | probable multi-dom |
| 45 | 91.5 | 5.3 | 13055 | 2 T16580 | hypothetical prote |

ALIGNMENTS

RESULT 1

T14782 hypothetical protein DKFP586B0621.1 - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Sep-2000

C/Accession: T14782

R/Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, August 1999

A/Reference number: Z18184

A/Accession: T14782

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-219 <OTT>

A/Cross-references: EMBL:AL110261

A/Experimental source: adult uterus; clone DKFP586B0621

C/Genetics:

A/Note: DKFP586B0621.1

C/Superfamily: complement C1q carboxyl-terminal homology

| Query Match | 12.3% | Score 213 | DB 2 | Length 219 |
|-----------------------|---|-------------------|---------------|------------|
| Best Local Similarity | 34.9% | Pred. No. 4.3e-10 | | |
| Matches | 66 | Conservative 18 | Mismatches 83 | Indels 22 |
| Gaps | | | | 6 |
| QY | 138 GAPHYALGAPGATFSG---YLVDADAD---ADAPARGPPAPP-----EPRSAFSAARTR 185 | | | |
| DB | 33 GAP-----GAPGEKGEGRGRLPGPRGDPGRGEAGPAGTGPAEGCSVPPRSAFSAKRS 88 | | | |
| QY | 186 SLVGSADAGPGRHQPLAFDTEFVNIGGDFDAAGVFRCLPGAYFFSFTLGLPKRLTSLV 245 | | | |
| DB | 89 SRV-----PPPSDAPLPFDRVLVNEQGHYDAVTGKTCQVPGVYFA-VHATVYRASLQF 142 | | | |
| QY | 246 KLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRRGDAVWLSHDHGCGAYSNHGKTYT 305 | | | |
| DB | 143 DLVKNGESIASFPQFFGWPFPKPSLSGGAMVRLPEPDQVWVQGVGDYIGIYASIKTST 202 | | | |
| QY | 306 FSGFLVYPD 314 | | | |
| DB | 203 FSGFLVYSD 211 | | | |

RESULT 2

JC4708 gelatin-binding 28K protein precursor - human

N/Alternate names: adipose specific collagen-like factor

C/Species: Homo sapiens (man)

C/Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 20-Sep-1999

C/Accession: JC4708; JC4944

R/Maeda, K.; Okubo, K.; Shimomura, I.; Funahashi, T.; Matsuzawa, Y.; Matsubara, K.

Biochem. Biophys. Res. Commun. 221, 286-289, 1996

A/Title: cDNA cloning and expression of a novel adipose specific collagen-like factor,

A/Reference number: JC4708; MUID:96224171; PMID:8619847

A/Accession: JC4708

A;Molecule type: mRNA
A;Residues: 1-244 <MAE>
A;Cross-references: DDBJ:D45371; NID:G871886; PIDN:BAA08227.1; PID:G871887
A;Experimental source: adipose tissue
R;Nakano, Y.; Tobe, T.; Choi-Miura, N.H.; Mazda, T.; Tomita, M.
J. Biochem. 120, 803-812, 1996
A;Title: Isolation and characterization of GBP28, a novel gelatin-binding protein purified from sheep erythrocytes
A;Reference number: JG4944; MUID:97103474; PMID:8947845
A;Accession: JG4944
A;Molecule type: protein
A;Residues: 19-38;93-100;101-112;135-149;173-178 <NAK>
C;Comment: This protein is an endogenous factor that binds with a collagen-like domain.
C;Genetics:
A;Gene: apM1
C;Superfamily: unassigned collagens; complement C1q carboxyl-terminal homology
C;Keywords: adipose tissue; glycoprotein; hydroxyproline
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-244/Product: gelatin-binding 28kDa protein #status experimental <MAT>
F;42-107/Region: collagen-like
F;114-241/Domain: complement C1q carboxyl-terminal homology <CIQ>
F;95/Modified site: 4-hydroxyproline (Pro) #status experimental
F;230/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 11.4%; Score 196; DB 2; Length 244;
Best Local Similarity 29.7%; Pred. No. 1.2e-08;
Matches 54; Conservative 31; Mismatches 61; Indels 36; Gaps 8;
QY 7 GLIGP-----AACWALGP-----TPPGSSSELRAFSARTPLEGSEMAV 48
DB 72 GLIGPKDIGETGVPGAEGPRGPGIQRKGEPEGAYVYRSAPSVGLEITYVT-IPNMP 130
QY 49 TFDKVVNIGGDFDVATGQFRCRVPGAYFFSFTAGKAPH-----KSLVMLVRNDEVQA 103
DB 131 RFTKIFVQNNHYDGTGKFFHCNIPGUYFAY-----HITVYMKDVKVSLEFK-KDKAML 183
QY 104 LAFDEQRPGARARASASAMLOLDYGDPTVWLRHGHAPH---VALGAPGATFSGYLTVYAD 159
DB 184 FTYDQYQENNVNQ-ASGSVLHLHLEVGQVWLQVYGEGERNGLYADNDNDSTFTGFLVYHD 242
QY 160 AD 161
DB 243 TN 244

RESULT 3
B48150
hibernation-related protein HP-25 precursor - Siberian chipmunk
C;Species: Eutamias sibiricus (Siberian chipmunk)
C;Date: 16-Feb-1994 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: B48150; B41752
R;Takamatsu, N.; Ohba, K.; Kondo, J.; Kondo, N.; Shiba, T.
Mol. Cell. Biol. 13, 1516-1521, 1993
A;Title: Hibernation-associated gene regulation of plasma proteins with a collagen-like
A;Reference number: A48150; MUID:93180798; PMID:8441393
A;Accession: B48150
A;Molecule type: mRNA; protein
A;Residues: 1-215 <TAK>
A;Cross-references: GB:D12975; NID:G287469; PIDN:BAA02352.1; PID:G287470
A;Note: the source is designated as Tamias asiaticus in Genbank entry TMSHP25, release 1
A;Note: sequence extracted from NCBI backbone (NCBIN:125946, NCBIP:125947)
R;Kondo, N.; Kondo, J.
J. Biol. Chem. 267, 473-478, 1992
A;Title: Identification of novel blood proteins specific for mammalian hibernation.
A;Reference number: A41752; MUID:92112696; PMID:1730610
A;Accession: B41752
A;Status: preliminary
A;Molecule type: protein
A;Residues: 29-62;84-130;172-183;187-192;201-215 <KON>
C;Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom
C;Keywords: glycoprotein; hibernation; plasma
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-215/Product: hibernation-related protein HP-25 #status experimental <MAT>
F;91-214/Domain: complement C1q carboxyl-terminal homology <CIQ>

F;167/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 10.7%; Score 184; DB 2; Length 215;
Best Local Similarity 30.4%; Pred. No. 9.6e-08;
Matches 62; Conservative 22; Mismatches 52; Indels 68; Gaps 11;
QY 138 GAPHYALGAPGATFSGYLVYADADADAPAPGPAPP-----EPRS 177
DB 49 GIPGFP-GAPGAL-----GPPGPPGVPGIPGPGPPGDYKCSSRPS 90
QY 178 AFSARTSLVSGDAGPGRHQPLAFDTEFVNICGDFDAAGVFRCLPGAYFFSFTLGK 237
DB 91 AFAVKL-----SERPEP-FQPIVFKEALYNGEHFNMATGEFSCVLPGVYNGFDI-R 142
QY 238 LPRKTLVKLMKNRDEVQAMIVDDGASRRREMOSQ-----SYMALRRGDVAVLLSH 289
DB 143 LFQSSVKIRLM--RDGIQV-----REKEAQANDSYKHAMGSVIMALGKDKVWLESK 192
QY 290 DHDGYGAYSNHG-KYITPFGFLVY 312
DB 193 LK--GTESEKGIHIVFFGYLLY 213

RESULT 4
C1HUQB
complement subcomponent C1q chain B precursor (validated) - human
N;Alternate names: complement subcomponent C1q beta chain
C;Species: Homo sapiens (man)
C;Date: 22-May-1981 #sequence_revision 31-May-1996 #text_change 08-Dec-2000
C;Accession: B23422; A23422; B50304; A90301; B90315; A03206
R;Reid, K.B.M.
Biochem. J. 231, 729-735, 1985
A;Title: Molecular cloning and characterization of the complementary DNA and gene codin
A;Reference number: A23422; MUID:86076906; PMID:3000358
A;Accession: B23422
A;Molecule type: DNA
A;Residues: 'HS', 1-32 <REI>
A;Note: the authors translated the codon AGT for the second position as Arg; they were
A;Accession: A23422
A;Molecule type: mRNA
A;Residues: 28-253 <REI>
A;Cross-references: EMBL:X03084
A;Note: the authors translated the codon ACA for residue 46 as Ile
R;Reid, K.B.M.
Biochem. J. 179, 367-371, 1979
A;Title: Complete amino acid sequences of the three collagen-like regions present in su
A;Reference number: A90304; MUID:80020137; PMID:486087
A;Accession: B90304
A;Molecule type: protein
A;Residues: 'E', 29-84, 'D', 86-99, 'P', 101-135 <RE5>
R;Reid, K.B.M.; Thompson, E.O.P.
Biochem. J. 173, 863-868, 1978
A;Title: Amino acid sequence of the N-terminal 108 amino acid residues of the B chain o
A;Reference number: A90301; MUID:79041552; PMID:708376
A;Accession: A90301
A;Molecule type: protein
A;Residues: 28-99, 'P', 101-195 <RE3>
R;Reid, K.B.M.; Gagnon, J.; Frampton, J.
Biochem. J. 203, 559-569, 1982
A;Title: Completion of the amino acid sequences of the A and B chains of subcomponent C
A;Reference number: A90315; MUID:82283890; PMID:6981411
A;Accession: B90315
A;Molecule type: protein
A;Residues: 136-253 <RE4>
A;Note: 176-Glx may also be present
C;Comment: The first component of complement is a calcium-dependent complex of the thre
ivation of C1r (enzyme), C1s (proenzyme), and the other eight components of complement
C;Comment: The C1q subcomponent is composed of nine subunits, six of which are disulfid
(see PIR:C1HUQC) chain. Equimolar amounts of the A, B, and C chains are found after red
C;Genetics:
A;Gene: GDB:C1QB
A;Cross-references: GDB:119043; OMIM:120570
A;Map position: 1p36.3-1p34.1

C;Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom
C;Keywords: complement pathway; glycoprotein; heterodimer; hydroxylysine; hydroxyproline
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-253/Product: complement subcomponent C1q chain B #status experimental <MAT>
F;33-116/Domain: collagenous, triple helix <COL>
F;123-249/Domain: complement C1q carboxyl-terminal homology <C1Q>
F;28/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F;31/Disulfide bonds: interchain (to chain A-26) #status experimental
F;35,38,41,53,56,65,83,86,101,104,107/Modified site: 4-hydroxyproline (Pro) #status expe
F;59,62,77,92,98,110/Modified site: 5-hydroxylysine (Lys) #status experimental
F;59,62,98,110/Binding site: carbohydrate (Lys) (covalent) #status experimental

Query Match 10.5%; Score 182; DB 1; Length 253;
Best Local Similarity 29.1%; Pred. No. 1.7e-07;
Matches 48; Conservative 32; Mismatches 61; Indels 24; Gaps 5;

QY 165 PARGPAPPEP-----RSAFSAARTSLVGSADGPPRHQPLAFDTEFVNIG 211
DB 97 PKGGPAGPAGPAPKGESGDYKATQKIAFSATRTINV-----PLRRDQTRFDHVTNMN 150
QY 212 GDFDAAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMT-YDGDASRRREM 270
DB 151 NNYPEPRSGKFTCKVFGLYFTYHASS--RGNLCVNLMRGRERAOQKVTFCDYAYNTFQVT 208
QY 271 SQSVMLALRRGDVAVMLLSHDHDGYGAYSNHGKYITFSGFLVYPDL 315
DB 209 TGGMVLKLEGGENVFLQATDKN--SLLMEGANSIFSGLFLFPDM 251

RESULT 5

S49158
complement protein C1q beta chain precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 20-Aug-1999
C;Accession: S49158
R;Schwaeble, W.; Petry, F.; Loos, M.
submitted to the EMBL Data Library, March 1993
A;Description: cDNA cloning and expression of the mRNA encoding the B-chain of rat C1q.
A;Reference number: S49158
A;Accession: S49158
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-253 <SCH>
A;Cross-references: EMBL:X71127; NID:g510191; PIDN:CA50440.1; PID:g510192
C;Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom
F;121-249/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 10.2%; Score 176.5; DB 2; Length 253;
Best Local Similarity 28.2%; Pred. No. 4.8e-07;
Matches 50; Conservative 40; Mismatches 66; Indels 21; Gaps 8;

QY 2 LPLLGLGPAACWALGP--TPGP-----GSSELR-----AFSAARTPLEGTSEMA 47
DB 80 IPGIPGKVGPKG--PVGPKAGPGRPGRPGKSGSDYKATQKVAFAALRTVNSALRPNQ 137
QY 48 VTFDKVYVNIIGDFDVATGQFRCRVPGAYFFSFTAGKAPHKSLSVMLVNRDEVQ-ALAF 106
DB 138 IRFEKVTITVNDNYEPRSGKFTCKVFGLYFTYHASSRGNLCVNIVGRDRDRMQKVLTF 197
QY 107 DEQRRPGARRAASQSAMLQLDYGDVTWMLRLHGAAPHYALGAPGAT--FSGYLVAADAD 161
DB 198 CDYAOQ-NTFQVTTGGVVLKLEQEEVHLQATD-KNSLLGVEGANSITGFLFLFPDM 252

RESULT 6

S23298
collagen alpha 1(VIII) chain - chicken
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S23298
R;Ninomiya, Y.; Castagnola, P.; Gerecke, D.; Gordon, M.K.; Jacenko, O.; LuValle, P.; Mc
maguchi, N.; Olsen, B.R.
in Extracellular Matrix Genes, Sandell L.J. and Boyd C.D., eds., pp.79-114, Academic Pre

A;Title: The molecular biology of collagens with short triple-helical domains.
A;Reference number: S22243
A;Accession: S23298
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-744 <NIN>
C;Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
F;617-743/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 10.1%; Score 175; DB 1; Length 744;
Best Local Similarity 25.3%; Pred. No. 2.4e-06;
Matches 56; Conservative 28; Mismatches 77; Indels 60; Gaps 8;

QY 136 LHGAPHYALGAPGATFSGYLVYADADADAPA-RGPPAPPEPR----- 177
DB 540 LHGPP-----GKPGA-----LGPGQPGLPGPPPGPPGPPAVMPPTPAQGEYLPDMG 589
QY 178 -----AFSAARTSLVGSADGPG-----PRHQPLAFDTEFVNIG 212
DB 590 LGIDGVKTPHAYAAK-----GKNGGPAYEMPAFTAELTAPFPVGAPIKFDRLLYNGRQ 644
QY 213 DFDAAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRRREMOSQ 272
DB 645 NYNPQTGIFTCEVPGVYFYFAHYH-HCKGGNVWVALFKNNEPVM-YTYDEYKKGFLDQASG 702
QY 273 SVMLALRRGDVAVMLLSHDHDGYGAYSNHGKYITFSGFLVYP 313
DB 703 SAVLLLRPGDRVFLQNPSEQAAGLYAGQYVHSSFSGYLLYP 743

RESULT 7

A34246
collagen alpha 1(VIII) chain precursor - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A34246
R;Yamaguchi, N.; Benya, P.D.; van der Rest, M.; Ninomiya, Y.
J. Biol. Chem. 264, 16022-16029, 1989
A;Title: The cloning and sequencing of alpha1(VIII) collagen cDNAs demonstrate that typ
omains similar to those of type X collagen.
A;Reference number: A34246; MUID:89380199; PMID:2476437
A;Accession: A34246
A;Molecule type: mRNA
A;Residues: 1-744 <YAM>
A;Cross-references: GB:J05042; NID:g164895; PIDN:AAA31204.1; PID:g164896
C;Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-744/Product: collagen alpha 1(VIII) chain #status predicted <MAT>
F;21-117/Region: amino-terminal nonhelical
F;118-571/Region: interrupted helical
F;572-744/Region: carboxyl-terminal nonhelical
F;617-743/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 10.0%; Score 173; DB 1; Length 744;
Best Local Similarity 25.3%; Pred. No. 3.4e-06;
Matches 56; Conservative 27; Mismatches 78; Indels 60; Gaps 8;

QY 136 LHGAPHYALGAPGATFSGYLVYADADADAPA-RGPPAPPEPR----- 177
DB 540 LHGPP-----GKPGA-----LGPGQPGLPGPPPGPPGPPAVMPPTPAQGEYLPDMG 589
QY 178 -----AFSAARTSLVGSADGPG-----PRHQPLAFDTEFVNIG 212
DB 590 LGIDGVKTPHAYAAK-----GKNGGPAYEMPAFTAELTAPFPVGAPIKFDRLLYNGRQ 644
QY 213 DFDAAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRRREMOSQ 272
DB 645 NYNPQTGIFTCEVPGVYFYFAHYH-HCKGGNVWVALFKNNEPVM-YTYDEYKKGFLDQASG 702
QY 273 SVMLALRRGDVAVMLLSHDHDGYGAYSNHGKYITFSGFLVYP 313
DB 703 SAVLLLRPGDRVFLQNPSEQAAGLYAGQYVHSSFSGYLLYP 743

```

RESULT 8
C1HUQC
complement subcomponent C1q chain C precursor - human
N;Alternate names: complement subcomponent C1q gamma chain
C;Species: Homo sapiens (man)
C;Date: 22-May-1981 #sequence_revision 31-May-1996 #text_change 22-May-1998
C/Accession: S14351; A03207
R;Seller, G.C.; Blake, D.J.; Reid, K.B.M.
Biochem. J. 274, 481-490, 1991
A;Title: Characterization and organization of the genes encoding the A-, B- and C-chains
A;Reference number: S14350; MUID:91174759; PMID:1706597
A/Accession: S14351
A;Status: not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-245 <SEL>
R;Reid, K.B.M.
Biochem. J. 179, 367-371, 1979
A;Title: Complete amino acid sequences of the three collagen-like regions present in sub
A;Reference number: A90304; MUID:80020137; PMID:486087
A/Accession: A03207
A/Molecule type: protein
A/Residues: 29-56,'P',58-65,'K',67-71,'P',73-83,'K',85-86,'D',88-89,'N',91-122 <RET>
C;Comment: The first component of complement is a calcium-dependent complex of the three
ivation of C1r (enzyme), C1s (proenzyme), and the other eight components of complement.
C;Comment: The C1q subcomponent is composed of nine subunits, six of which are disulfide
dimers of the C chain. Equimolar amounts of the A, B, and C chains are found after redu
C;Genetics:
A;Gene: GDB:C1QC
A;Cross-references: GDB:128132; OMIM:120575
A;Map position: 1p36.3-1p34.1
A;introns: 60/3
C;Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom
C;Keywords: complement pathway; glycoprotein; homodimer; hydroxylysine; hydroxyproline;
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-245/Product: complement subcomponent C1q chain B #status predicted <MAT>
F;31-114/Domain: collagenous, triple helix <COL>
F;121-244/Domain: complement C1q carboxyl-terminal homology <C1Q>
F;32/Disulfide bonds: interchain #status experimental
F;36,39,42,45,54,63,81,93,96,99,105/Modified site: 4-hydroxyproline (Pro) #status experi
F;57,72,75/Modified site: 5-hydroxylysine (Lys) #status experimental
F;75/Binding site: carbohydrate (Lys) (covalent) #status experimental

Query Match          10.0%; Score 172; DB 1; Length 245;
Best Local Similarity 23.8%; Pred. No. 1.1e-06;
Matches 67; Conservative 33; Mismatches 84; Indels 98; Gaps 13;

QY      86 PHKSLSVMLVRNRDEVQALAFDEQRRPGARRAASQSAMQLDYGDTVWLRLHGAPHYALG 145
Db       9 PHLGKLLLL-----LLLLALRGQANTGC-----YG-----IPGMPGLP-G 43

QY      146 APGATFSGY--LVYADADADAP-----RGPAPP----- 173
Db       44 APGK--DGYDGLPGPKGEPGIPALPGIRGPKGQKGEPGLPGHPKNGMPGMPGVGP 101

QY      174 -----EP-----RSAFSAARTSLVSGSDAGPGPRHQP-----LAFDTEFVNIGG 212
Db      102 MGIPGEPEEGRGYKQKQFSVFVTR-----QTHQPPAPNSLIRFNAVLTPQG 149

QY      213 DFDAAAGVFRCLPGAYFFSFTLGKLPKRTL SVKLMKNRDEVQAMIVDDGASRRREMOSQ 272
Db      150 DYDTSTGKFTCKVPGLYYFVYHAS---HTANLCVLLYRSGVAVVTFCGHTSKTNQVNSG 205

QY      273 SVMLALRRGDAVWLLSHDHDGYGAYSNHGKYITFSGFLVYPPD 314
Db      206 GVLRLRYGEEVWLAVNDY--YDMVGIQGSDSVFSGFLLFPD 245

RESULT 9
S15435
collagen alpha 1(VIII) chain precursor - human
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Aug-1999

```

C;Accession: S15435
R;Muragaki, Y.; Mattei, M.G.; Yamaguchi, N.; Olsen, B.R.; Ninomiya, Y.
Eur. J. Biochem. 197, 615-622, 1991
A;Title: The complete primary structure of the human alpha-1(VIII) chain and assignment
A;Reference number: S15435; MUID:91231001; PMID:2029894
A;Accession: S15435
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-744 <MUR>
A;Cross-references: EMBL:X57527; NID:g30081; PIDN:CAA40748.1; PID:g30082
C;Genetics:
A;Gene: GDB:COL8A1
A;Cross-references: GDB:128104; OMIM:120251
A;Map position: 3q11.1-3q13.2
C;Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-744/Product: collagen alpha 1(VIII) chain #status predicted <MAT>
F;21-117/Region: amino-terminal nonhelical
F;118-571/Region: interrupted helical
F;572-744/Region: carboxyl-terminal nonhelical
F;617-743/Domain: complement C1q carboxyl-terminal homology <C1Q>

| Query Match | 9.8%; | Score 170; | DB 2; | Length 744; |
|-----------------------|---|------------------|------------|-------------|
| Best Local Similarity | 25.9%; | Pred. No. 6e-06; | | |
| Matches 56; | Conservative 27; | Mismatches 83; | Indels 50; | Gaps 9 |
| QY | 136 LHGAPHYALGAPGATFSGYLVYADADADAPA-RGPPAPPEPRS----- | 177 | | |
| Db | 540 LHGPP----GKPGA-----LGPGQGPGUGPPGPPGPPAPVMPPTPPQGEYLPDMG | 589 | | |
| QY | 178 -AFSAARTSLVGS DAGP--GPRHQLAFDTE---FVNIG-----DFDA | 217 | | |
| Db | 590 LGIDGYRPHATGAKKKGNGSPAYEMBAFTALTAFFPPVGGPVKFNKLLYNGRONYNPQ | 649 | | |
| QY | 218 AGVFRCLPGAYFFSFTLGLPKRTLSTVKLMKNRDEVQAMITYDGASRRRNOQSQSVMLA | 277 | | |
| Db | 650 TGIFTCEVPGVYYFAYHV-HCKGGNVVALFKNNEPVM-ITYDEYKKGFLDQASGSAVLL | 707 | | |
| QY | 278 LRRGDAVWLLSHDHDGYGAYSNHGKIYTFSGFLVYP | 313 | | |
| Db | 708 LRPGRVFLQMPSEQAAGLYAGQYVHSSFSGYLLYP | 743 | | |

RESULT 10

S23297
collagen alpha 1(X) chain precursor - chicken

N:Alternate names: type X collagen
C:Species: Gallus gallus (chicken)
C:Date: 07-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 13-Aug-1999
C:Accession: S23297; A31896; S65594; S77711; I50218
R:Ninomiya, Y.; Castagnola, P.; Gerecke, D.; Gordon, M.K.; Jacenko, O.; LuValle, P.; Maguchi, N.; Olsen, B.R.
in Extracellular Matrix Genes, Sandell L.J. and Boyd C.D., eds., pp.79-114, Academic Press, 1992

A:Title: The molecular biology of collagens with short triple-helical domains.
A:Reference number: S22243

A:Accession: S23297
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-674 <NIN>

R:LuValle, P.; Ninomiya, Y.; Rosenblum, N.D.; Olsen, B.R.
J. Biol. Chem. 263, 18378-18385, 1988

A:Title: The type X collagen gene. Intron sequences split the 5'-untranslated region and
A:Reference number: A31896; MUID:89054019; PMID:2461368

A:Accession: A31896
A:Molecule type: mRNA
A:Residues: 1-75 <LUV>

R:Ninomiya, Y.; Gordon, M.; van der Rest, M.; Schmid, T.; Linsenmayer, T.; Olsen, B.R.
J. Biol. Chem. 261, 5041-5050, 1986

A:Title: The developmentally regulated type X collagen gene contains a long open reading
A:Reference number: I50218; MUID:86168227; PMID:3082876

A:Accession: S65594
A:Molecule type: DNA
A:Residues: 'T', '9', 'D', '11-12', 'EDQMKLYLFTM', '30-31', 'TCKSGRAFTTYMILQNMADLVSSH', '48-89', 'L'

```

629, 'PQAVLSLISWRTIKCGSSCOIQNPMVSIPLMFIILSQVSYYLLKSNIPLTMS' <NNI>
A;/Cross-references: EMBL:M13496; NID:g211699; PIDN:AAA48736.1; PID:g211700
A;/Accession: S77711
A;/Molecule type: protein
A;/Residues: 104-112,'X',114-117;453-466 <NIN2>
C;/Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
C;/Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer; hydroxyproline;
F;1-18/Domain: signal sequence #status predicted <SIG>
F;547-673/Domain: complement C1q carboxyl-terminal homology <CIQ>
F;453,456/Modified site: hydroxyproline (Pro) #status experimental
F;611/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          9.8%; Score 169.5; DB 2; Length 674;
Best Local Similarity 25.3%; Pred. No. 5.8e-06;
Matches 58; Conservative 21; Mismatches 75; Indels 75; Gaps 7;

QY      145 GAPGATFSGYLVADADAAPA-----RGPAPPEPR----- 176
        |||||           |:|||
Db       454 GAGP-----AKGEAGAPGLPGPATGTGLRGPMGPDPGPKGNSGEPGLPGPPG 504
        |||||           |:|||

QY      177 -----SAPSAARTSLVG-----SDAGPGGRHQPL 201
        |||||           |:|||
Db       505 PPGPPGSTIPEGYVKGSRELSGSMFKAGANQALTGMPSAFTVLISKAYPGAT-VPI 563
        |||||           |:|||

QY      202 AFDTEFNIGGFDAAGVFRCLPGAYFFSFLLGLRKTL SVKL MKNRDEVQAMTYDD 261
        |||||           |:|||
Db       564 KFDKILYNRQQHYDPRGTGIFTCRIPLGYFSYHV-HAKGTNWVALYKNQSPVM-YTYDE 621
        |||||           |:|||

QY      262 GASRRREMOSQSVMLALNRGDVAWVLLSHDH DGYGVYNSNHGKYITFSGFL 310
        :||::|||::|||::|||::|||::|||
Db       622 YQKGYLDAQSGSAVIDLMENDQVWLQLPNSESNGLYSSEYVHSSFSGFL 670

RESULT 11
collagen alpha 1(X) chain precursor - mouse
S31216
C;/Species: Mus musculus (house mouse)
C;/Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text change 13-Aug-1999
C;/Accession: S31216; S28807; S22215; S30127; I48299; S26397; S31830
R;/Kong, R.Y.C.; Kwan, K.M.; Lau, E.T.; Thomas, J.T.; Boot-Handford, R.P.; Grant, M.E.;
Eur. J. Biochem. 213, 99-111, 1993
A;/Title: Introns-exon structure, alternative use of promoter and expression of the mouse
A;/Reference number: S31216; MUID:93238750; PMID:8477738
A;/Accession: S31216
A;/Molecule type: DNA
A;/Residues: 1-680 <KON>
A;/Cross-references: EMBL:Z21610; NID:g49793; PIDN:CAA79736.1; PID:g49794
R;/Elima, K.; Eroola, I.; Rosati, R.; Metsaeranta, M.; Garofalo, S.; Peraelae, M.; de Croc
Biochem. J. 289, 247-253, 1993
A;/Title: The mouse collagen X gene: complete nucleotide sequence, exon structure and exp
A;/Reference number: S28807; MUID:93143676; PMID:8424763
A;/Accession: S28807
A;/Molecule type: DNA
A;/Residues: 1-285,'A','287-680 <ELI>
A;/Cross-references: EMBL:X67348; NID:g50480; PIDN:CAA47763.1; PID:g50481
R;/Elima, K.; Metsaeranta, M.; Kalilo, J.; Peraelae, M.; Eroola, I.; Garofalo, S.; de Croc
Biochim. Biophys. Acta 1130, 78-80, 1992
A;/Title: Specific hybridization probes for mouse alpha-2(IX) and alpha-1(X) collagen mRN
A;/Reference number: S22215; MUID:92182017; PMID:1543751
A;/Accession: S22215
A;/Status: preliminary
A;/Molecule type: mRNA
A;/Residues: 385-450,'K',452-627 <ELA>
A;/Cross-references: EMBL:X63013; NID:g49795; PIDN:CAA44741.1; PID:g49796
R;/Appe, S.S.; Olsen, B.R.
Matrix 13, 165-179, 1993
A;/Title: Characterization of the mouse type X collagen gene.
A;/Reference number: S30127; MUID:93261348; PMID:8492743
A;/Accession: S30127
A;/Status: preliminary
A;/Molecule type: mRNA
A;/Residues: 1-12,'F',14-26,'S',28-247,'L',249-285,'A',287-305,'F',307-416,'S',418-499,'L
R;/Appe, S.S.; Seldin, M.F.; Hayashi, M.; Olsen, B.R.
```

```

Eur. J. Biochem. 206, 217-224, 1992
A;Title: Cloning of the human and mouse type X collagen genes and mapping of the mouse
A;Reference number: 148299; MUID:92267014; PMID:1587271
A;Accession: 148299
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 52-247, 'L', 249-285, 'A', 287-305, 'F', 307-416, 'S', 418-499, 'L', 501-566, 'C', 568,
A;Cross-references: EMBL:X5121; NID:g50482; PIDN:CAA46237.1; PID:g667031
J. Sumners, T.A.; Irwin, M.H.; Mayne, R.; Balian, G.
J. Biol. Chem. 263, 581-587, 1988
A;Title: Monoclonal antibodies to type X collagen. Biosynthetic studies using an antibody
A;Accession: S26397
A;Molecule type: protein
A;Residues: 'SDGYFSQ', 24-26, 'KQ' <SUM>
C;Genetics:
A;Gene: Col10a-1
A;Map position: 10
A;Introns: 51/3
C;Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
C;Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-680/Product: collagen alpha 1(X) chain #status predicted <MAT>
F;553-679/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match          9.8%; Score 168.5; DB 2; Length 680;
Best Local Similarity 25.7%; Pred. No. 7.1e-06;
Matches 56; Conservative 23; Mismatches 82; Indels 57; Gaps 6;

QY      145 GAGGATFSGYLVYADADADAPARGPPAPPEBSAFAAARTSLVSGDAGPGRHQ----- 199
Db      470 GNPGA--PGPAGIATKGLNGPT-GPPGPGR--GHSGEPLGPDPGPGRPGQAVMPD 523

QY      200 -----PLAFDTEFVNIGDFF 215
Db      524 GFIKAGQRPRLSGMPLVSANHGVTGMPVSAFTVILSKAYPAVGAPIPDEILYNRQHYD 583

QY      216 AAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRRREMOSQSV 275
Db      584 PRSGIFTCKIPGIYYFSYHV-HVKGTHVWGLYKNGTPTM-YTYDEYSKGYLDQASGSAI 641

QY      276 LALRRGDAVWLISHDHDGYGAYSNHGKYITFSGFLVYP 313
Db      642 MELTENDQVWLQLPNAESNGLYSSEYVHSSFSGLVAP 679

RESULT 12
S19018
complement subcomponent C1q chain A precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Aug-1999
C;Accession: S19018
R;Petty, F.; Reid, K.B.M.; Loos, M.
J. Immunol. 147, 3988-3993, 1991
A;Title: Gene expression of the A- and B-chain of mouse C1q in different tissues and th
A;Reference number: S19018; MUID:92043789; PMID:1940381
A;Accession: S19018
A;Molecule type: mRNA
A;Residues: 1-245 <PET>
A;Cross-references: EMBL:X58861; NID:g50226; PIDN:CAA41664.1; PID:g50227
C;Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal ho
F;116-243/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match          9.7%; Score 168; DB 2; Length 245;
Best Local Similarity 31.0%; Pred. No. 2.3e-06;
Matches 53; Conservative 16; Mismatches 70; Indels 32; Gaps 4;

QY      6  LGLGPACVAGLPTPGSSSELRSAAARTPLEGTSEMAVFDKYYVNIIGDFDVAT 65
Db      91  LGDSGPQGLKGVKGNPNIRDPRAFAIRONPM--TLGNVYTFDKVLTNOESPYNHT 148

QY      66  GQRCRVPGAYFFSF-----TAGKAPHKSLSLVLRNDEVQALAFDEQR 111

```


Db 149 GRFICAVPGFYFNQVISKMDLCLFIKSSSGGQPRDLSFSNTNKGLEQVLA----- 202
QY 112 PGARRAASQSAMLDYDGTWLRHLGHAPH--YALGAPGATFSGYLVYADA 160
Db 203 -----GGTVLQLRRGDEWIEKDPAGRIYQGTREADSIFSGFLIFPSA 245

RESULT 13

CGHUID

collagen alpha 1(X) chain precursor - human

N;Alternate names: procollagen alpha 1(X) chain

C;Species: Homo sapiens (man)

C;Date: 22-Nov-1993 #sequence revision 03-Nov-1995 #text change 22-Jun-1999

C;Accession: S26396; S30086; S15826; S18249; A43901; I51870; S21856

R;Reichenberger, E.; Beier, F.; LuValle, P.; Olsen, B.R.; von der Mark, K.; Bertling, W.

FEBS Lett. 311, 305-310, 1992

A;Title: Genomic organization and full-length cDNA sequence of human collagen X.

A;Reference number: S26396; MUID:93012005; PMID:1397333

A;Accession: S26396

A;Molecule type: DNA

A;Residues: 1-680 <REI>

A;Cross-references: EMBL:X68952; EMBL:X72578; EMBL:X72579; EMBL:X72580; GB:S47714; GB:S4

R;Appe, S.S.

submitted to the EMBL Data Library, March 1992

A;Reference number: S30085

A;Accession: S30086

A;Molecule type: DNA

A;Residues: 'TIPFYGVWCWVCL', 52-680 <APT>

A;Cross-references: EMBL:X65120; NID:g23129

A;Note: the initial difference is probably due to translation of an intronic sequence

R;Appe, S.; Mattei, M.G.; Olsen, B.R.

FEBS Lett. 282, 393-396, 1991

A;Title: Cloning of human alpha-1(X) collagen DNA and localization of the COL10A1 gene

A;Reference number: S15826; MUID:91243838; PMID:2037056

A;Accession: S15826

A;Molecule type: DNA

A;Residues: 561-647, 'G', 649-666 <AP2>

A;Cross-references: EMBL:X58879; NID:g30013; PIDN:CAA41686.1; PID:g30014

R;Thomas, J.T.; Cresswell, C.J.; Rash, B.; Nicolai, H.; Jones, T.; Solomon, E.; Grant, M

Biochem. J. 280, 617-623, 1991

A;Title: The human collagen X gene. Complete primary translated sequence and chromosomal

A;Reference number: S18249; MUID:92109659; PMID:1764025

A;Accession: S18249

A;Molecule type: DNA

A;Residues: 1-26, 'T', 28-680 <THO>

A;Cross-references: EMBL:X60382; NID:g30094; PIDN:CAA42933.1; PID:g30095

A;Note: the sequence from Fig. 3 is inconsistent with that from Fig. 2 in having 532-Ala

R;Reichenberger, E.; Aigner, T.; von der Mark, K.; Stoss, H.; Bertling, W.

Dev. Biol. 148, 562-572, 1991

A;Title: In situ hybridization studies on the expression of type X collagen in fetal hum

A;Reference number: A43901; MUID:92077285; PMID:1743401

A;Accession: A43901

A;Molecule type: mRNA

A;Residues: 547-656 <RE2>

A;Cross-references: GB:M74050; GB:D57494; NID:g339884; PIDN:AAA61221.1; PID:g553796

A;Note: sequence extracted from NCBI backbone (NCBIN:69012, NCBI:P:69014)

R;Wallis, G.A.; Rash, B.; Sweetman, W.A.; Thomas, J.T.; Super, M.; Evans, G.; Grant, M.E

Am. J. Hum. Genet. 54, 169-178, 1994

A;Title: Amino acid substitutions of conserved residues in the carboxyl-terminal domain

pe Schmid.

A;Reference number: I51870; MUID:94136476; PMID:8304336

A;Accession: I51870

A;Status: translated from GB/EMBL/DBD

A;Molecule type: mRNA

A;Residues: 520-597, 'D', 599-680 <WAL>

A;Cross-references: GB:S68531; NID:g545180; PIDN:AAC60615.1; PID:g545181

A;Note: mutant sequence from patient with metaphyseal chondrodysplasia type Schmid

C;Comment: Prolines and lysines at the third position of the tripeptide repeating unit

ed and subsequently O-glycosylated.

C;Genetics:

A;Gene: GDB:COL10A1

A;Cross-references: GDB:128635; OMIM:120110

A;Map position: 6q21-6q22

A;introns: 52/1

A;Note: a defect in this gene may cause Schmid metaphyseal chondrodysplasia

C;Complex: type X collagen may be a homotrimer

C;Function:

A;Description: structural component of extracellular fibrous polymer specifically and c

be important for skeletogenesis

C;Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology

C;Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer; hydroxyllysine;

F;1-18/Domain: signal sequence #status predicted <SIG>

F;19-680/Product: collagen alpha 1(X) chain #status predicted <MAT>

F;19-56/Domain: amino-terminal nonhelical #status predicted <NC2>

F;57-519/Region: interrupted helical

F;520-680/Domain: amino-terminal nonhelical #status predicted <NC1>

F;553-679/Domain: complement C1q carboxyl-terminal homology <C1Q>

F;617/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

9.6%; Score 166.5; DB 1; Length 680;

Best Local Similarity 25.7%; Pred. No. 1e-05;

Matches 56; Conservative 22; Mismatches 83; Indels 57; Gaps 6;

QY 145 GAPGATFSGYLVYADADADAPARGPAPPEPRSAFSAARTSLVSGDAGPGRHQ----- 199

Db 470 GSRGP--PGPAGIATKGLNGPT-GPPGPGR--GHSGEPLGPBPBPBGQAVMPE 523

QY 200 -----PLAPDTEFVNIGDFD 215

Db 524 GFIKAGQRPSTSGTFLVSANQGTGMPVSATVILSKAYPAIGTPIPFDLKLYNRQHYD 583

QY 216 AAGVFRCLPGAYFFSFTLGLPRKTLSTVKLMKNRDEVQAMITYDDGASRRREMOSQSV 275

Db 584 PRTGIFTCCQIPGIYFFSYHV-HVKGTHVMVGLYKNGTPVM-YTYDEYTKGYLDQASGSAI 641

QY 276 LALRRGDAVWLLSHDHGCGAYSNHGKIYTFSGFLVYP 313

Db 642 IDLTENDQVWLQLPNAESNGLYSSEYVHSSFSGLVAP 679

RESULT 14

collagen alpha 1(VIII) chain - mouse

C;Species: Mus musculus (house mouse)

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000

C;Accession: S23779

R;Muragaki, Y.; Shiota, C.; Inoue, M.; Ooshima, A.; Olsen, B.R.; Ninomiya, Y.

Eur. J. Biochem. 207, 895-902, 1992

A;Title: Alpha-1(VIII)-collagen gene transcripts encode a short-chain collagen polypept

A;Reference number: S23779; MUID:92362626; PMID:1499564

A;Accession: S23779

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-743 <MUR>

A;Cross-references: EMBL:X66976; NID:g50493; PIDN:CAA47387.1; PID:g1359953

C;Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology

F;616-742/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match

9.6%; Score 165.5; DB 1; Length 743;

Best Local Similarity 25.8%; Pred. No. 1.4e-05;

Matches 56; Conservative 23; Mismatches 85; Indels 53; Gaps 6;

QY 136 LHGAPHYALGAPGATFSGYLVYADADADAPARGPAPPEPRSAFSAARTSLVGS---DA 192

Db 540 LHGPP---GKPGA-----LGPQGPGLPBPBPBPBPBPVMPPTSPQGEYLPDM 587

QY 193 GPG-----PRHQPLAFDTEFVNIGDFDA 216

Db 588 GLGIDGVKTPHAYAGKKKGHGEPAVEMPAFTAEITVPPEVGAIVKFDKLLYNGRQYNP 647

QY 217 AAGVFRCLPGAYFFSFTLGLPRKTLSTVKLMKNRDEVQAMITYDDGASRRREMOSQSV 276

Db 648 QTGITFCEVPGVYFYFAHYV-HCKGQNVWVALFKN-NEPMYTYTYDEYKKGFLDQASGSAVL 705

QY 277 ALRRGDAVWLLSHDHGCGAYSNHGKIYTFSGFLVYP 313

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:09 ; Search time 11.8204 Seconds

(without alignments)
1308.910 Million cell updates/sec

Title: US-10-085-167-2

Perfect score: 1726

Sequence: 1 MLLPLLLGLLPACWALGPT.....LVYPDLAPAPPGIGASELL 329

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------|--------------------|
| 1 | 1726 | 100.0 | 329 | 1 CQT4_HUMAN | Q9bxj3 homo sapien |
| 2 | 215 | 12.5 | 281 | 1 CQT1_HUMAN | Q9bxj1 homo sapien |
| 3 | 213 | 12.3 | 243 | 1 CQT5_HUMAN | Q9bxj0 homo sapien |
| 4 | 208 | 12.1 | 285 | 1 CQT2_HUMAN | Q9bxj5 homo sapien |
| 5 | 200 | 11.6 | 278 | 1 CQT6_HUMAN | Q9bx19 homo sapien |
| 6 | 199 | 11.5 | 258 | 1 C1RF_MOUSE | O88992 mus musculu |
| 7 | 196 | 11.4 | 244 | 1 APM1_HUMAN | Q15848 homo sapien |
| 8 | 196 | 11.4 | 247 | 1 APM1_MOUSE | Q60994 mus musculu |
| 9 | 192 | 11.1 | 258 | 1 C1RF_HUMAN | Q75973 homo sapien |
| 10 | 186.5 | 10.8 | 255 | 1 GLIC_MOUSE | Q9esn4 mus musculu |
| 11 | 185.5 | 10.7 | 289 | 1 CQT7_HUMAN | Q9bxj2 homo sapien |
| 12 | 184 | 10.7 | 215 | 1 HP25_TAMSI | Q06576 tamias sibi |
| 13 | 182 | 10.5 | 251 | 1 C1QB_HUMAN | P02746 homo sapien |
| 14 | 176.5 | 10.2 | 253 | 1 C1QB_RAT | P31721 rattus norv |
| 15 | 173.5 | 10.1 | 419 | 1 COLE_LEPMA | P14282 oryctolagus |
| 16 | 173 | 10.0 | 744 | 1 C1A18_RABIT | P98085 lepomis mac |
| 17 | 170 | 9.8 | 245 | 1 C1QC_HUMAN | P02747 homo sapien |
| 18 | 168.5 | 9.8 | 680 | 1 C1A1_MOUSE | Q05306 mus musculu |
| 19 | 168 | 9.7 | 245 | 1 C1QA_MOUSE | P98086 mus musculu |
| 20 | 167.5 | 9.7 | 674 | 1 C1A1_CHICK | P08125 gallus gall |
| 21 | 167 | 9.7 | 744 | 1 C1A18_HUMAN | P27658 homo sapien |
| 22 | 166.5 | 9.6 | 253 | 1 C1QB_MOUSE | P14106 mus musculu |
| 23 | 166.5 | 9.6 | 680 | 1 C1A1_HUMAN | Q03692 homo sapien |
| 24 | 166.5 | 9.6 | 743 | 1 C1A18_MOUSE | Q00780 mus musculu |
| 25 | 163.5 | 9.5 | 246 | 1 CQT3_HUMAN | Q9bxj4 homo sapien |
| 26 | 163.5 | 9.5 | 674 | 1 C1A1_BOVIN | P23206 bos tauris |
| 27 | 163 | 9.4 | 224 | 1 CERL_RAT | P98087 rattus norv |
| 28 | 163 | 9.4 | 508 | 1 OTO1_ONCKE | P83371 oncorhynch |
| 29 | 160.5 | 9.3 | 246 | 1 C1QC_MOUSE | Q02105 mus musculu |
| 30 | 158 | 9.2 | 193 | 1 CERB_HUMAN | P23435 homo sapien |
| 31 | 156.5 | 9.1 | 193 | 1 CERB_MOUSE | Q9r171 mus musculu |
| 32 | 151.5 | 8.8 | 245 | 1 C1QA_HUMAN | P02745 homo sapien |
| 33 | 150 | 8.7 | 196 | 1 HP20_TAMSI | Q06575 tamias sibi |

| | | | | | |
|----|-------|-----|------|--------------|--------------------|
| 34 | 147.5 | 8.5 | 215 | 1 HP27_TAMSI | Q06577 tamias sibi |
| 35 | 147 | 8.5 | 1228 | 1 ECM_HUMAN | Q13201 homo sapien |
| 36 | 139.5 | 8.1 | 635 | 1 CA28_HUMAN | P25067 homo sapien |
| 37 | 139 | 8.1 | 201 | 1 CERL_HUMAN | Q9ntu7 homo sapien |
| 38 | 135.5 | 7.9 | 170 | 1 CA28_MOUSE | P25318 mus musculu |
| 39 | 98.5 | 5.7 | 265 | 1 Y176_HUMAN | Q14681 homo sapien |
| 40 | 97.5 | 5.6 | 476 | 1 LEU2_STRCO | O86534 streptomyce |
| 41 | 96.5 | 5.6 | 497 | 1 FXD2_HUMAN | O60548 homo sapien |
| 42 | 95.5 | 5.5 | 493 | 1 YKAY_CAEEL | P34258 caenorhabdi |
| 43 | 95 | 5.5 | 495 | 1 DHAL_ECOLI | P23883 escherichia |
| 44 | 94 | 5.4 | 1045 | 1 GUNB_CELFI | P26225 cellulomona |
| 45 | 93.5 | 5.4 | 686 | 1 VGLH_PRVN3 | Q00660 pseudorabie |

ALIGNMENTS

| RESULT 1 | ID | QOT4_HUMAN | STANDARD; | PRT; | 329 AA. |
|--|--|---|-----------|---------------------------------------|---------|
| AC | Q9BXJ3; | | | | |
| DT | 28-FEB-2003 (Rel. 41, Created) | | | | |
| DT | 28-FEB-2003 (Rel. 41, Last sequence update) | | | | |
| DT | 28-FEB-2003 (Rel. 41, Last annotation update) | | | | |
| DE | Complement-clq tumor necrosis factor-related protein 4 precursor. | | | | |
| GN | C1QTNF4 OR CTRP4. | | | | |
| OS | Homo sapiens (Human). | | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | | |
| OX | NCBI_TaxID=9606; | | | | |
| RN | [1] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RA | Holloway J.L., Lok S.; | | | | |
| RT | "Homo sapiens complement-clq tumor necrosis factor-related protein."; | | | | |
| RL | Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases. | | | | |
| CC | -1- SIMILARITY: Contains 2 C1Q domains. | | | | |
| CC | | | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | | | |
| CC | or send an email to license@isb-sib.ch). | | | | |
| CC | | | | | |
| DR | EMBL; AF329838; AAK17962.1; - | | | | |
| DR | Genew; HGNC:14346; C1QTNF4. | | | | |
| DR | InterPro; IPR001073; C1q. | | | | |
| DR | Pfam; PF00386; C1q; 2. | | | | |
| DR | SMART; SM00110; C1Q; 2. | | | | |
| DR | PROSITE; PS01113; C1Q; 2. | | | | |
| KW | Repeat; signal. | | | | |
| FT | SIGNAL | 1 | 16 | POTENTIAL. | |
| FT | CHAIN | 17 | 329 | COMPLEMENT-C1Q TUMOR NECROSIS FACTOR- | |
| FT | | | | RELATED PROTEIN 4. | |
| FT | DOMAIN | 23 | 159 | C1Q 1. | |
| FT | DOMAIN | 170 | 314 | C1Q 2. | |
| SQ | SEQUENCE | 329 AA; | 35265 MW; | 331C7DBF26036915 CRC64; | |
| Query Match | | | | | |
| Best Local Similarity 100.0%; Score 1726; DB 1; length 329; | | | | | |
| Matches 329; Conservativity 0; Mismatches 0; Indels 0; Gaps 0; | | | | | |
| QY | 1 | MLPLLLGLLPACWALGPTPGGSSELSAFAARTTPLEGTSEMAVTFDKYVYVIGGD | 60 | | |
| Db | 1 | MLPLLLGLLPACWALGPTPGGSSELSAFAARTTPLEGTSEMAVTFDKYVYVIGGD | 60 | | |
| QY | 61 | FDVATGQRCRVPGAYFFSFTAGKAPHKSLVLRNDEVQALAFDEQRRGARRASQ | 120 | | |
| Db | 61 | FDVATGQRCRVPGAYFFSFTAGKAPHKSLVLRNDEVQALAFDEQRRGARRASQ | 120 | | |
| QY | 121 | SAMQLDYGDTVWRLHGAPHYALGAPGATFSGYLVYADADADAPARGPAPBPERSAFS | 180 | | |

Db 121 SAMQLDYGDTWRLRHGAPHYALGAPGATFSGYLVYADADADAPAPGPAPPEPRSAFS 180
Qy 181 AARTSLVGSADAGPGRHOPPLAFDTEFVNIGGDFDAAAGVFRCLPGAYFFSFTLGKLP 240
Db 181 AARTSLVGSADAGPGRHOPPLAFDTEFVNIGGDFDAAAGVFRCLPGAYFFSFTLGKLP 240
Qy 241 KTLVKLMKNRDEVQAMITYDDGASRRRMOQSVMALRRGDVWILSHDHGCGAYSNH 300
Db 241 KTLVKLMKNRDEVQAMITYDDGASRRRMOQSVMALRRGDVWILSHDHGCGAYSNH 300
Qy 301 GKYITFSGFLVYPDLAPAPGLGASELL 329
Db 301 GKYITFSGFLVYPDLAPAPGLGASELL 329

RESULT 2
COT1 HUMAN STANDARD; PRT; 281 AA.
ID COT1_HUMAN
AC Q9BXJ1; Q96NF2; Q9GZR4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement-clq tumor necrosis factor-related protein 1 precursor
DE (G protein coupled receptor interacting protein) (GIP).
GN C1QTNF1 OR CTRP1.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Innamorati G., Le Gouill C., Whang I., Birnbaumer M.;
RT "GIP, a putative GPCR interacting protein.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

RN [2]
RP SEQUENCE FROM N.A.
RA Sheppard P.O.;
RT "Homo sapiens complement-clq tumor necrosis factor-related protein.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [4]
RP SEQUENCE OF 83-281 FROM N.A.
RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA Hotuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J.,
RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Magatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K., Isegai T.;

RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C1Q domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AJ272138; CAC20425.1; ALT_INIT.
DR EMBL; AF232905; AAG44303.1; ALT_INIT.
DR EMBL; AF329840; AAK17964.1; -.
DR EMBL; BC021553; AAH21553.1; -.
DR EMBL; AK055541; BAB70947.1; -.
DR Genew; HGNC:14324; C1QTNF1.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q, FALSE_NEG.
FW SIGNAL 1 25
FT CHAIN 26 281
FT FT
FT DOMAIN 99 140
FT FT
FT CONFLICT 22 22
FT CONFLICT 241 241
SQ SEQUENCE 281 AA; 31743 MW; 49E248CB88ACFB7C CRC64;
L -> P (IN REF. 1).
R -> Q (IN REF. 1).
POTENTIAL.
COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
RELATED PROTEIN 1.
COLLAGEN-LIKE.
C1Q.
L -> P (IN REF. 1).
R -> Q (IN REF. 1).
SEQUENCE 281 AA; 31743 MW; 49E248CB88ACFB7C CRC64;

Query Match 12.5%; Score 215; DB 1; Length 281;
Best Local Similarity 30.2%; Pred. No. 2.7e-10;
Matches 77; Conservative 26; Mismatches 86; Indels 66; Gaps 11;

Qy 69 RCRVPGAYFFSFTAGKAPHKSLVNLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDY 128
Db 75 RCCDPGTSMPYATA--VPQINITILKGEKDR-----GDRG-----LQGY 113
Qy 129 GDT--VWLRHGA---HYALGAPGATFSGYLVYADADADAPAPGPAPPEPRSAFSAAR 183
Db 114 GKTGSAGARGHTGPXQKSGMGAPGERCKSH--YA-----AFSVGR 152
Qy 184 TRSLVGSADAGPGRH-----QPLAFDTEFVNIGGDFDAAAGVFRCLPGAYFFSFTLGKL 238
Db 153 KK-----PMHSNHYQTIVIFDTEFVNLYDHFMFTGKFCYVPGLYFFSLNVHTW 202
Qy 239 PRKTLVKLMKNRDEVQAMITYDDGASRRRMOQSVMALRRGDVWILSHDHGCGAYS 298
Db 203 NQKETYLMKNNEEVILFAQVG--DRSIMQSIMLELREQDQVWVRLYGERENALIF 260
Qy 299 NH--GKYITFSGFLV 311
Db 261 SEELDTYITFSGYLV 275

RESULT 3
COT5 HUMAN STANDARD; PRT; 243 AA.
ID COT5_HUMAN
AC Q9BXJ0; Q9UFY4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Complement-clq tumor necrosis factor-related protein 5 precursor.
GN C1QTNF5 OR CTRP5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;

```
RN [1]
RP SEQUENCE FROM N.A.
RA Sheppard P.O., Humes J.M.;
RT "Homo sapiens complement-clq tumor necrosis factor-related protein.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 25-243 FROM N.A.
RC TISSUE=Uterus;
RA Ottenwaelder B., Obermaier B., Mewes H.-W., Gassenhuber J.,
RA Wiemann S.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 ClQ domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
DR EMBL; AF329841; AAK17965.1; -
DR EMBL; AL110261; CAB53702.1; -
DR PIR; T14782; T14782.
DR Genew; HGNC:14344; C1QTNF5.
DR InterPro; IPR001073; Clq.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; Clq; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; FALSE_NEG.
DR KMW
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 16 243 COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT RELATED PROTEIN 5.
FT DOMAIN 30 95 COLLAGEN-LIKE.
FT DOMAIN 97 243 C1Q.
SQ SEQUENCE 243 AA; 25298 MW; 7CCDA65CDA7EB784 CRC64;

Query Match 12.3%; Score 213; DB 1; Length 243;
Best Local Similarity 34.9%; Pred. No. 3.3e-10;
Matches 66; Conservative 18; Mismatches 83; Indels 22; Gaps 6;

OY 138 GAPHYALGAPGATFSG---YLVDADAD---ADAPARGPAPP-----EPRSAFSAATR 185
Db 57 GAP-----GAPGKEGEGRPGLPGPRGDPGRGEAGPAGPTGPAEGCSVPPRSAFSKRSE 112
OY 186 SLVGSDDAGPGRHQPLAFDTEFVNIGDFDAAGVRCRLPGAYFFSFTLGKLPKTLSTV 245
Db 113 SRV-----PPPSDAPLPFDRVLVNEQGHYDAVTGKTCQVPGVYFA-VHATVYRASLQF 166
OY 246 KLKMKRDEVQAMTYDDGASRRREMOSQSVMLALRGDAVWLSDHHDGYGAYSNHGKYIT 305
Db 167 DLVKNGESIASFFQFGGWPKPASLSGAMVRLPEPDQVWVGVDYIGIYASIKTDST 226
OY 306 FSGFLVYPD 314
Db 227 FSGFLVYSD 235

RESULT 4
COT2_HUMAN STANDARD; PRT; 285 AA.
AC Q9BXU5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement-clq tumor necrosis factor-related protein 2 precursor.
GN C1QTNF2 OR CTRP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Piddington C.S., Bishop P.;
RT "Homo sapiens complement-clq tumor necrosis factor-related protein.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Garninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 ClQ domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
DR EMBL; AF329836; AAK17960.1; -
DR EMBL; BC011699; AAH11699.1; -
DR Genew; HGNC:14325; C1QTNF2.
DR InterPro; IPR001073; Clq.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; Clq; 1.
DR Pfam; PF01391; Collagen; 2.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
DR KMW
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 16 285 COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT RELATED PROTEIN 2.
FT DOMAIN 40 141 COLLAGEN-LIKE.
FT DOMAIN 143 285 C1Q.
SQ SEQUENCE 285 AA; 29952 MW; 7E31FF9868D4EDFA CRC64;

Query Match 12.1%; Score 208; DB 1; Length 285;
Best Local Similarity 33.3%; Pred. No. 1e-09;
Matches 55; Conservative 31; Mismatches 49; Indels 30; Gaps 8;

OY 7 GLGPAACWALGPTPGGSSELRSFAFAA--RTTPLEGTSEMAVFDKYVNVNIGDFDVA 64
Db 136 GLPGPCSC-----GSGHTKSAFSVAVTXSYPRE--RLPIKDKILMNEGGHYNAS 183
OY 65 TGQFRCRVPAGAYFFSFTAGKAPHKSLVMLVRNDEVQALAFDEQRRPGARRASOSAMT 124
Db 184 SGKFCVCGVGEIYYFTYDITLA-NKHLAIGLVHN-GQYRIRTFD--ANTGNHDVASGSTIL 239
OY 125 QLDYGDYTWLRHLGA-----PHYALGAPGATFSGYLVDAD 161
```


Db 240 ALKOGDEWVLQIFYSSEONGLFYDPYWT----DSLFTGFLIYADQD 280

RESULT 5
COT6_HUMAN STANDARD; PRT; 278 AA.
AC Q9BXI9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement-c1q tumor necrosis factor-related protein 6 precursor.
GN C1QTNF6 OR CTRP6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Piddington C.S., Sheppard P.O.;
RT "Homo sapiens complement-c1q tumor necrosis factor-related protein.",
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loqueilano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C1Q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF329842; AAK17966.1; -.
DR EMBL; BC020551; AAH20551.1; -.
DR Genew; HGNC:14343; C1QTNF6.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR PROSITE; PS01113; C1Q; FALSE_NEG.
KW Collagen; Signal.
FT SIGNAL 1 46
FT CHAIN 47 278
FT DOMAIN 97 138
FT DOMAIN 139 278
FT CARBOHYD 91 91
POTENTIAL.
COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
RELATED PROTEIN 6.
COLLAGEN-LIKE.
C1Q.
N-LINKED (GLCNAC. .) (POTENTIAL).

FT CONFLICT 21 21 G -> V (IN REF. 2).
SQ SEQUENCE 278 AA; 30861 MM; 27A82CA863F23D47 CRC64;
Query Match 11.6%; Score 200; DB 1; Length 278;
Best Local Similarity 31.3%; Pred. No. 4.3e-09;
Matches 87; Conservative 21; Mismatches 106; Indels 64; Gaps 12;

QY 48 VTFDKVYVNIIGGDEVDATGQFRC-----RVPGAYFFSFTAGKAPHKSLSVMLVRNDEVQ 102
Db 46 LTFDRA-----VASGCQRCCDSEDPDPAHVSSASSSGRPHALPET----- 86
QY 103 ALAFDEQRFGARRAASQASMLQLDYGDYVWLRHGAAPHY-----ALGAPGATFSGYLVY 157
Db 87 -----RPYNITILKGDKDPGPM---GLPGYMGREGPGQGFQGS----- 125
QY 158 ADADADAPARGPP-APPEPR-SAFSAARTSLVGS DAGPRHQPLAFPTFEFVNIIGDFD 215
Db 126 ---KGDKGEMGSPGAPCQKRFPAFVSVKRTAL-----HSGEDFQTLFERYVNLDCFPD 177
QY 216 AAAGVFRCLPGAYFFSFTLGLPRKTLVXLMKNRDEVQAMVYDDGASRRREMOSQSV 275
Db 178 MATGQFAAPLRGIYFSLVHWSMNYKETVYHIMNQKE--AVILYAQPSERSIMOSQSV 235
QY 276 LALRGDAVWLSHDHGYGA-YSN-HGKYTFESGFLV 311
Db 236 LDLAYGDRVWVRLFRKQRENAIYSNDFDTYTFSGHLI 273

RESULT 6
C1RF_MOUSE STANDARD; PRT; 258 AA.
AC 08892;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C1q-related factor precursor.
GN C1QRF OR CRF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99097006; PubMed=9878755;
RA Berube N.G., Swanson X.H., Bertram M.J., Kittle J.D., Didenko V.,
RA Baskin D.S., Smith J.R., Pereira-Smith O.M.;
RT "Cloning and characterization of CRF, a novel C1q-related factor,
RT expressed in areas of the brain involved in motor function."
RL Brain Res. Mol. Brain Res. 63:233-240 (1999).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAINSTEM. MORE ABUNDANT IN AREAS
CC OF THE NERVOUS SYSTEM INVOLVED IN MOTOR FUNCTION, SUCH AS THE
CC PURKINJE CELLS OF THE CEREBELLUM, THE ACCESSORY OLIVARY NUCLEUS,
CC THE PONS AND THE RED NUCLEUS.
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C1Q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF095155; AAC64187.1; -.
DR MGD; MGI:1344400; C1QRF.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

| | | | | |
|----|-----|--|---|---------------------------------------|
| | DR | EMBL; D45371; | BAA08227.1; | -. |
| | DR | EMBL; AB012165; | BAA86716.1; | -. |
| | DR | EMBL; AB012164; | BAA86716.1; | JOINED. |
| | DR | EMBL; AJ131460; | CAB52413.1; | -. |
| | DR | EMBL; AJ131461; | CAB52413.1; | JOINED. |
| | DR | PIR; JC4708; | JC4708. | |
| | DR | MIM; 605441; | -- | |
| | DR | GO; GO:0006091; | P:energy pathways; TAS. | |
| | DR | InterPro; IPR01073; | C1q. | |
| | DR | InterPro; IPR000087; | Collagen. | |
| | DR | pfam; PF00386; | C1q; 1. | |
| | DR | pfam; PF01391; | collagen; 1. | |
| | DR | PRINTS; PR00007; | COMPLEMENTC1Q. | |
| | DR | PRODOM; PD000007; | C1q_helix; 1. | |
| | DR | SMART; SM00110; | C1Q; 1. | |
| | KM | PROSITE; PS01113; | C1Q; 1. | |
| | KW | Hormone; Collagen; Signal; Repeat; Hydroxylation; Plasma; | | |
| | KV | Polymorphism; Disease mutation; Obesity; Diabetes mellitus. | | |
| | FT | SIGNAL | 1 14 POTENTIAL. | |
| | FT | CHAIN | 15 244 ADIPONECTIN. | |
| | FT | DOMAIN | 42 107 COLLAGEN-LIKE. | |
| | FT | DISELID | 108 244 C1Q. | |
| | FT | MOD_RES | 36 36 INTERCHAIN (BY SIMILARITY). | |
| | FT | MOD_RES | 44 44 HYDROXYLATION (BY SIMILARITY). | |
| | FT | MOD_RES | 47 47 HYDROXYLATION (BY SIMILARITY). | |
| | FT | MOD_RES | 53 53 HYDROXYLATION (BY SIMILARITY). | |
| | FT | MOD_RES | 62 62 HYDROXYLATION (BY SIMILARITY). | |
| | FT | MOD_RES | 71 71 HYDROXYLATION (BY SIMILARITY). | |
| | FT | MOD_RES | 76 76 HYDROXYLATION (BY SIMILARITY). | |
| | FT | MOD_RES | 86 86 HYDROXYLATION (BY SIMILARITY). | |
| | FT | MOD_RES | 95 95 HYDROXYLATION (BY SIMILARITY). | |
| | FT | MOD_RES | 104 104 HYDROXYLATION (BY SIMILARITY). | |
| | FT | VARIANT | 84 G -> R. | |
| | FT | | /FTid=VAR_013273. | |
| | FT | VARIANT | 112 112 R -> C (in adiponectin deficiency). | |
| | FT | | /FTid=VAR_013274. | |
| | FT | VARIANT | 117 117 V -> M. | |
| | FT | | /FTid=VAR_013275. | |
| | FT | VARIANT | .164 .164 I -> T. | |
| | FT | | /FTid=VAR_013276. | |
| | FT | VARIANT | 221 221 R -> S. | |
| | FT | | /FTid=VAR_013277. | |
| | FT | VARIANT | 241 241 H -> P. | |
| | FT | | /FTid=VAR_013278. | |
| | SO | SEQUENCE | 244 AA; 26414 MW; 64D8C6CI204B1018 CRC64; | |
| | | Query Match | Best Local Similarity | 11.4%; Score 196; DB 1; Length 244; |
| | | Matches | 54; Conservative | 31; Mismatches 61; Indels 36; Gaps 8, |
| OY | 7 | G LUGP-----AACWALGP-----TPPGSSSELRAFSARTTPLEGTSEMAV | 48 | |
| DB | 72 | GLIGPKGDIGETGVPGAEGPPFGIQRKGEPEGAGAYVRSAFVSGLETTYVT-IPNMP | 130 | |
| OY | 49 | TEDKVYNIGGDFDVATGGFRGVPGAFFSFSTAKAPH----KSLSVMLVRNRDEVQA | 103 | |
| DB | 131 | RFTKIIFYNQNNHYDGSTGKFHCNIPLGYFYA-----HITVMKDVKYSLEK-KDKAM | 183 | |
| OY | 104 | LAFDEQRRPGARRASQSAMLQLDYGDTWMRLHGAPH---YAIGAFCATESGYLVAD | 159 | |
| DB | 184 | FYYDYQEENNVDQ-ASSSVLLHLLEVGDQVWLQVYGEGERNGLYADNDNSTFTGFLLYHD | 242 | |
| OY | 160 | AD 161 | | |
| DB | 243 | TN 244 | | |

RESULT 8
APM1 MOUSE

ID APM1_MOUSE STANDARD; PRT; 247 AA.
 AC Q60994; Q62400; Q9DC68;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Adiponectin precursor (30 kDa adipocyte complement-related protein)
 DE (ACRP30) (Adipocyte specific protein AdipoQ).
 GN APM1 OR ACRP30 OR ADIPOQ.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Adipocyte;
 RX MEDLINE=96070757; PubMed=7592907;
 RA Scherer P.E., Williams S., Fogliano M., Baldini G., Lodish H.F.;
 RT "A novel serum protein similar to C1q, produced exclusively in
 adipocytes.";
 RL J. Biol. Chem. 270:26746-26749 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fibroblast;
 RX MEDLINE=96209999; PubMed=8631877;
 RA Hu E., Liang P., Spiegelman B.M.;
 RT "AdipoQ is a novel adipose-specific gene dysregulated in obesity.";
 RL J. Biol. Chem. 271:10697-10703 (1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX PubMed=1162643;
 RA Das K., Lin Y., Widen E., Zhang Y., Scherer P.E.;
 RT "Chromosomal localization, expression pattern, and promoter analysis
 of the mouse gene encoding adipocyte-specific secretory protein
 Acrp30.";
 RL Biochem. Biophys. Res. Commun. 280:1120-1129 (2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Heart;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Giusti C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staahl F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Wittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690 (2001).
 RN [5]
 RP FUNCTION.
 RX MEDLINE=21372498; PubMed=11479627;
 RA Yamauchi T., Kamon J., Waki H., Teruchi Y., Kubota N., Hara K.,
 RA Mori Y., Ide T., Murakami K., Tsuboyama-Kasaoka N., Ezaki O.,
 RA Akenuma Y., Gavrilova O., Vinson C., Reitman M.L., Kagechika H.,
 RA Shudo K., Yoda M., Nakano Y., Tobe K., Nagai R., Kimura S., Tomita M.,
 RA Froguel P., Kadowaki T.;
 RT "The fat-derived hormone adiponectin reverses insulin resistance
 associated with both lipodystrophy and obesity.";
 RL Nat. Med. 7:941-946 (2001).
 RN [6]
 RP FUNCTION.
 RX MEDLINE=21372499; PubMed=11479628;

| | |
|----|------------------------------------|
| RP | FUNCTION. |
| RX | MEDLINE=21372499; Pubmed=11479628; |

RA Berg A.H., Combs T.P., Du X., Brownlee M., Scherer P.E.;
RT "The adipocyte-secreted protein Acrp30 enhances hepatic insulin
action.";
RL Nat. Med. 7:947-953(2001).
CC -|- FUNCTION: IMPORTANT NEGATIVE REGULATOR IN HEMATOPOIESIS AND IMMUNE
CC SYSTEMS; MAY BE INVOLVED IN ENDING INFLAMMATORY RESPONSES THROUGH
CC ITS INHIBITORY FUNCTIONS. INHIBITS ENDOTHELIAL NF-KAPPA-B SIGNALING
CC THROUGH A CAMP-DEPENDENT PATHWAY. INHIBITS TNF-ALPHA-INDUCED
CC EXPRESSION OF ENDOTHELIAL ADHESION MOLECULES. INVOLVED IN THE
CC CONTROL OF FAT METABOLISM AND INSULIN SENSITIVITY.
CC -|- SUBUNIT: Homooligomer.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- TISSUE SPECIFICITY: Synthesized exclusively by adipocytes and
CC secreted into plasma.
CC -|- INDUCTION: DURING HORMONE-INDUCED ADIPOSE DIFFERENTIATION AND
CC ACTIVATED BY INSULIN.
CC -|- SIMILARITY: Contains 1 collagenous domain.
CC -|- SIMILARITY: Contains 1 C1q domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U37222; AAA80543.1; -.
DR EMBL; U49915; AAB06706.1; -.
DR EMBL; AF304466; AAK13417.1; -.
DR EMBL; AK003138; BAB22597.1; -.
DR PDB; 1C28; 07-SEP-99.
DR MGD; MGI:106675; Acrp30.
DR GO; GO:0005576; C:extracellular; IDA.
DR GO; GO:0005515; F:protein binding activity; IMP.
DR GO; GO:0006635; P:fatty acid beta-oxidation; IMP.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR ProDom; PD000007; C1q helix; 1.
DR SMART; SM00110; C1q; 1.
DR PROSITE; PS01113; C1q; 1.
KW Hormone; Collagen; Signal; Repeat; Hydroxylation; Plasma;
KW Polymorphism; 3D-structure.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 247 ADIPONECTIN.
FT DOMAIN 45 110 COLLAGEN-LIKE.
FT DOMAIN 111 247 C1Q.
FT DISULFID 39 39 INTERCHAIN (BY SIMILARITY).
FT MOD_RES 47 47 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 50 50 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 56 56 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 65 65 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 79 79 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 98 98 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 107 107 HYDROXYLATION (BY SIMILARITY).
FT VARIANT 113 113 M -> V.
FT CONFLICT 50 50 P -> S (IN REF. 2).
FT CONFLICT 74 74 A -> S (IN REF. 2).
FT CONFLICT 117 117 A -> G (IN REF. 2).
FT CONFLICT 148 148 G -> N (IN REF. 2).
FT CONFLICT 243 243 Y -> F (IN REF. 2).
SQ SEQUENCE 247 AA; 26841 MW; 137B687D873988C4 CRC64;

Query Match 11.4%; Score 196; DB 1; Length 247;
Best Local Similarity 31.9%; Pred. No. 7.8e-09;
Matches 58; Conservative 30; Mismatches 58; Indels 36; Gaps 11;

QY 7 GLLGP-----AACWALGP-----TPG-----PGSSE--LRSAFSARTTPLEGTSEMAV 48
DB 75 GLLGPKGETGVDVGMTGAEGPRGFGPTGPKRGEPGEALMYRSAFSVGLETRYI-VPNVPI 133

QY 49 TEDKVVYVIGGDFPVATGCFRCRVPGAYFFSFSTAGKAPH-----KSLSVMLVNRDEVQA 103
DB 134 RFTKIFYNQNHVDGSGTGKFCYCNIPGLYYFSY-----HITVYMKDVKVSIFK-KDKAVL 186
QY 104 LAFDEQRRPGRARRAASQSAMLQLDYDGTWLRHGG-APH---YALGAPGATPSGYLYAD 159
DB 187 FTYDQYQEKXNDQ-ASGSVLLHLEVGDVWLQVYGDGDHNGLYADNVNDSTFGRLLYHD 245
QY 160 AD 161
DB 246 TN 247

RESULT 9
C1RF_HUMAN STANDARD; PRT; 258 AA.
ID C1RF_HUMAN
AC 075973;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE C1q-related factor precursor.
GN C1QRF OR CRF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99097006; PubMed=9878755;
RA Berube N.G., Swanson X.H., Bertram M.J., Kittle J.D., Didenko V.,
RA Baskin D.S., Smith J.R., Pereira-Smith O.M.;
RT "Cloning and characterization of CRF, a novel C1q-related factor,
RT expressed in areas of the brain involved in motor function.";
RL Brain Res. Mol. Brain Res. 63:233-240(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA Nickerson D.A.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -|- TISSUE SPECIFICITY: EXPRESSED IN BRAINSTEM.
CC -|- SIMILARITY: Contains 1 collagenous domain.
CC -|- SIMILARITY: Contains 1 C1q domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial


```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF095154; AAC64186.1; -.
DR EMBL: AF410771; AAK95248.1; -.
DR EMBL: BC008798; AAH08798.1; -.
DR GO: 0007626; P: locomotory behavior; NAS.
DR InterPro: IPR001073; Clq.
DR InterPro: IPR000087; Collagen.
DR Pfam: PF00386; Clq; 1.
DR Pfam: PF01391; Collagen; 1.
DR PRINTS: PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Collagen; Signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 258 C1Q-RELATED FACTOR.
FT DOMAIN 67 115 COLLAGEN-LIKE.
FT DOMAIN 123 258 C1Q.
SQ SEQUENCE 258 AA; 26452 MW; 52C51CDF59CAE2BF CRC64;

Query Match 11.1%; Score 192; DB 1; Length 258;
Best Local Similarity 27.3%; Pred. No. 1.7e-08;
Matches 67; Conservative 26; Mismatches 88; Indels 64; Gaps 9;

QY 110 RRPGARRAASQSAMLQLDYCDTWLRLHGAPHYAL-----GAPGATFSGLVYADADADA.164
   ||||| : ||||| : |||||
Db 38 RGPAG-----ARTDGDALSEQSGAPPSTLVGSGKPGRT-----GKP 78

QY 165 PARGPAPPEPRSAFSAARTSLVGSADGPG-----PR-----197
   ||||| : ||||| : |||||
Db 79 GPPGPPGDPGPPGPPVGPGEKEGPKPPGPGLPAGGSGAISTATYTTVPRAFYAGLKN 138

QY 198 ----HQLAFDTEFVNIGDGFDAAGVFCRLPGAYFFSFTLGLPR---KTLGVKMK 249
   ::||| : ||||| : ||||| : ||||| : |||||
Db 139 PHEGYEVLKFDVVTNLGNVYDAASGKFTCNIPGTFFTYHV--LMRGDGTSMWADLCK 196

QY 250 NRDEVQAMLYDDGASRRREMOSQSVMLALRRGDAVWLSHDHGAYSNHGKYTFSGF 309
   |::| : : : ||||| : ||||| : |||||
Db 197 N-GQVRASAIQDADQNYDYASNSVILHLADAGDEVFIKLDGKAHGNSN--KYSTFSGF 253

QY 310 LVYPD 314
   ::|||
Db 254 ILYSD 258

RESULT 10
GLIC_MOUSE STANDARD; PRT; 255 AA.
ID_GLIC_MOUSE
AC Q9ESN4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gliacolin precursor (C1q-like protein).
GN C1QL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20428709; PubMed=10862616;
RA Koide T., Aso A., Yoriizu T., Nagata K.;
RT "Conformational requirements of collagenous peptides for recognition
RT by the chaperone protein HSP47."
RL J. Biol. Chem. 275:27957-27963(2000).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN GLIAL CELLS.
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C1Q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

[illegible]

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- SIMILARITY: Contains 1 collagenous domain.
CC -!- SIMILARITY: Contains 1 C1Q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF329839; AAK17963.1; -.
DR EMBL; BC022187; AAH22187.1; -.
DR Genew; HGNC:14342; C1QTNF7.
DR InterPro; IPR001073; C1Q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 2.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Collagen; Signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 289 COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT RELATED PROTEIN 7.
FT DOMAIN 38 139 COLLAGEN-LIKE.
FT DOMAIN 141 276 C1Q.
SQ SEQUENCE 289 AA; 30683 MW; A61609FF86D26946 CRC64;

Query Match 10.7%; Score 185.5; DB 1; Length 289;
Best Local Similarity 33.5%; Pred. No. 6.6e-08;
Matches 60; Conservative 22; Mismatches 60; Indels 37; Gaps 9;

QY 7 GLLGPACWALGPTGP-----GSSSLRSAPSAARTT-PLGEGSEM 46
Db 113 GEVGP-----IGP-PGPKGDRREGDPLPGVCRCGSIVLKSAFSGVITTSYPEE--RL 163

QY 47 AVTFDKVYVNIIGDFDVATGQRCRVPGAYFFSFATGKAPKSLVLMVRNDEVQALAF 106
Db 164 PIFPNKVLNEDGEHNPATKFIKAFPGIYFSYDITLA-NKHLAIGLVHN-GQYRIKTF 221

QY 107 DEQRRFGARRAASOSAMLQLDYGDVTWLRHLGAPHYAL---GAPGATFSGYLVYADAD 161
Db 222 D--ANTGNHDVASGSTVILQPEDEVWLEIFFTDQNGLFSDPGWADSLFSGFLLYVDTD 278

RESULT 12
HP25_TAMSI
ID HP25_TAMSI STANDARD; PRT; 215 AA.
AC 006576;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hibernation-associated plasma protein HP-25 precursor (Hibernator-
DE specific blood complex, 25 kDa subunit).
OS Tamias sibiricus (Siberian chipmunk) (Asian chipmunk).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
OC Tamias.
OX NCBI_TaxID=64680;
RN [1]

RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93180798; PubMed=8441393;
RA Takamatsu N., Ohba K., Kondo J., Kondo N., Shiba T.,
RT "Hibernation-associated gene regulation of plasma proteins with a
RT collagen-like domain in mammalian hibernators.";
RL Mol. Cell. Biol. 13:1516-1521(1993).
RN [2]
RP SEQUENCE OF 29-62; 84-130; 172-183; 187-192 AND 201-215.
RC TISSUE=Plasma;
RX MEDLINE=92112696; PubMed=1730610;
RA Kondo N., Kondo J.,
RT "Identification of novel blood proteins specific for mammalian
RT hibernation.";
RL J. Biol. Chem. 267:473-478(1992).
CC -!- FUNCTION: PLASMA PROTEINS HP-20, HP-25, HP-27 AND HP-55 FORM A
CC 140 kDa COMPLEX VIA DISULFIDE BONDS IN THE PLASMA AND ARE
CC HIBERNATION SPECIFIC.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
CC -!- DEVELOPMENTAL STAGE: THE PROTEIN COMPLEX DISAPPEARS FROM THE
CC PLASMA AT ONSET OF HIBERNATION AND REAPPEARS AS HIBERNATION
CC CEASES.
CC -!- SIMILARITY: Contains 1 collagenous domain.
CC -!- SIMILARITY: Contains 1 C1Q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D12975; BAA02352.1; -.
DR PIR; B48150; B48150.
DR InterPro; IPR001073; C1Q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Signal; Collagen; Glycoprotein; Plasma; Multigene family.
FT SIGNAL 1 28
FT CHAIN 29 215 HIBERNATION-ASSOCIATED PLASMA PROTEIN
FT HP-25.
FT DOMAIN 40 81 COLLAGEN-LIKE.
FT DOMAIN 83 215 C1Q.
FT CARBOHYD 167 167 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 215 AA; 22664 MW; AFE03206917EA530 CRC64;

Query Match 10.7%; Score 184; DB 1; Length 215;
Best Local Similarity 30.4%; Pred. No. 6.2e-08;
Matches 62; Conservative 22; Mismatches 52; Indels 68; Gaps 11;

QY 138 GAPHYALGAPGATFSGYLVYADADADAPARGPAPP-----EPRS 177
Db 49 GIPGFP-GAPGAL-----GPPGPPGVPGIPGPGPPGDVEKCSSRPKS 90

QY 178 AFSARTSLVSGDAGPGPRHQPPLAFDTFVNIGDFDAAGVFRCLPGAYFFSFTLGK 237
Db 91 AFAVKL-----SERPPEP-FQPIVFKEALYNQGHFMATGEFSCVLPGVYNGFDI-R 142

QY 238 LPRKTLVKLMKNRDEVQAMITYDDGASRRRMSQ-----SVMLALRGGDAVWLLSH 289
Db 143 LFQSSVKIRLM--RDGIQV-----REKEAQANDSYKHAMGSVIMLGKGDKWLLESK 192

QY 290 DHGQYAYSNG-KYITESGFLVY 312
Db 193 LK---GTSESEKITHIVFGYLLY 213

```

RESULT 13
C1QB_HUMAN STANDARD; PRT: 251 AA.
ID C1QB_HUMAN
AC P02746; Q96H17;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement C1q subcomponent, B chain precursor.
GN C1QB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86076906; PubMed=3000358;
RA Reid K.B.M.;
RT "Molecular cloning and characterization of the complementary DNA and
RT gene coding for the B-chain of subcomponent C1q of the human
RT complement system.";
RL Biochem. J. 231:729-735(1985).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 26-133.
RX MEDLINE=80020137; PubMed=486087;
RA Reid K.B.M.;
RT "Complete amino acid sequences of the three collagen-like regions
RT present in subcomponent C1q of the first component of human
RT complement.";
RL Biochem. J. 179:367-371(1979).
RN [4]
RP SEQUENCE OF 26-193.
RX MEDLINE=79041552; PubMed=708376;
RA Reid K.B.M., Thompson E.O.P.;
RT "Amino acid sequence of the N-terminal 108 amino acid residues of the
RT B chain of subcomponent C1q of the first component of human
RT complement.";
RL Biochem. J. 173:863-868(1978).
RN [5]
RP SEQUENCE OF 134-251.
RX MEDLINE=82283890; PubMed=6981411;
RA Reid K.B.M., Gagnon J., Frampton J.;
RT "Completion of the amino acid sequences of the A and B chains of
RT subcomponent C1q of the first component of human complement.";
RL Biochem. J. 203:559-569(1982).
RN [6]
RP SEQUENCE OF 224-251 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=85038855; PubMed=6208566;
FT Reid K.B.M., Bentley D.R., Wood K.J.;

```

```

RT "Cloning and characterization of the complementary DNA for the B
RT chain of normal human serum C1q.";
RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 306:345-354(1984).
RN [7]
RP REVIEW OF C1Q DEFICIENCY.
RX MEDLINE=98450587; PubMed=9777412;
RA Petry F.;
RT "Molecular basis of hereditary C1q deficiency.";
RL Immunobiology 199:286-294(1998).
CC -!- FUNCTION: C1Q ASSOCIATES WITH THE PROENZYMES C1R AND C1S TO YIELD
CC C1, THE FIRST COMPONENT OF THE SERUM COMPLEMENT SYSTEM. THE
CC COLLAGEN-LIKE REGIONS OF C1Q INTERACT WITH THE CA(2+)-DEPENDENT
CC C1R(2)C1S(2) PROENZYME COMPLEX, AND EFFICIENT ACTIVATION OF C1
CC TAKES PLACE ON INTERACTION OF THE GLOBULAR HEADS OF C1Q WITH THE
CC FC REGIONS OF IGG OR IGM ANTIBODY PRESENT IN IMMUNE COMPLEXES.
CC -!- SUBUNIT: C1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q, R
CC AND S IN THE MOLAR RATION OF 1:2:2. C1Q SUBCOMPONENT IS COMPOSED
CC OF NINE SUBUNITS, SIX OF WHICH ARE DISULFIDE-LINKED DIMERS OF THE
CC A AND B CHAINS, AND THREE OF WHICH ARE DISULFIDE-LINKED DIMERS OF
CC THE C CHAIN.
CC -!- PTM: O-linked glycans consist of Glc-Gal disaccharides bound to
CC the oxygen atom of post-translationally added hydroxyl groups.
CC -!- DISEASE: Defects in C1QB are a cause of C1q deficiency
CC [MIM:120570]. It is a rare genetic disorder which is associated
CC with recurrent infections and a high prevalence of lupus
CC erythematosus-like symptoms. It is characterized by a loss of
CC activation of the complement classical pathway.
CC -!- SIMILARITY: Contains 1 collagenous domain.
CC -!- SIMILARITY: Contains 1 C1Q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X03084; CAA26880.1; -
DR EMBL; BC008983; AA08983.1; ALT_INIT.
DR EMBL; M36278; AAC41692.1; -
DR Genew; HGNC:1242; C1QB.
DR MIM; 120570; -
DR GO; GO:0005602; C:complement component C1q complex; TAS.
DR GO; GO:0003811; F:complement activity; TAS.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PRO0007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Complement pathway; Plasma; Hydroxylation; Glycoprotein; Collagen;
KW Repeat; Signal; Disease mutation; Pyrrolidone carboxylic acid.
FT SIGNAL 1 25
FT CHAIN 26 251 COMPLEMENT C1Q SUBCOMPONENT, B CHAIN.
FT DOMAIN 29 112 COLLAGEN-LIKE.
FT DOMAIN 113 251 C1Q.
FT MOD_RES 26 26 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 29 29 INTERCHAIN (WITH C-26 IN CHAIN A).
FT MOD_RES 33 33 HYDROXYLATION.
FT MOD_RES 36 36 HYDROXYLATION.
FT MOD_RES 39 39 HYDROXYLATION.
FT MOD_RES 42 42 HYDROXYLATION.
FT MOD_RES 51 51 HYDROXYLATION.
FT MOD_RES 54 54 HYDROXYLATION.
FT MOD_RES 57 57 HYDROXYLATION.
FT CARBOHYD 57 57 O-LINKED (GAL. . .).
FT MOD_RES 60 60 HYDROXYLATION.
FT CARBOHYD 60 60 O-LINKED (GAL. . .).
FT MOD_RES 63 63 HYDROXYLATION.
FT MOD_RES 75 75 HYDROXYLATION.
FT MOD_RES 81 81 HYDROXYLATION.

```


[illegible]

```

CC      THE C CHAIN. IN ADDITION TO THE MAJOR A:B AND C:C DIMER BANDS,
CC      RAT, UNLIKE HUMAN C1Q, CONTAINED MINOR DIMER SPECIES.
CC      -1- SIMILARITY: Contains 1 collagenous domain.
CC      -1- SIMILARITY: Contains 1 C1q domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X71127; CAA50440.1; -.
DR      PIR; S49158; S49158.
DR      InterPro; IPR001073; C1q.
DR      InterPro; IPR000087; Collagen.
DR      Pfam; PF00386; C1q; 1.
DR      Pfam; PF01391; Collagen; 1.
DR      PRINTS; PR00007; COMPLEMENTC1Q.
DR      SMART; SM00110; C1Q; 1.
DR      PROSITE; PS01113; C1Q; 1.
KW      Complement pathway; Plasma; Hydroxylation; Glycoprotein; Collagen;
FT      Repeat; Signal.
FT      SIGNAL 1 25 BY SIMILARITY.
FT      CHAIN 26 253 COMPLEMENT C1Q SUBCOMPONENT, B CHAIN.
FT      DOMAIN 29 112 COLLAGEN-LIKE.
FT      DOMAIN 113 253 C1Q.
FT      DISULFID 29 29 INTERCHAIN (WITH C-26 IN CHAIN A).
FT      MOD_RES 33 33 HYDROXYLATION (BY SIMILARITY).
FT      MOD_RES 36 36 HYDROXYLATION (BY SIMILARITY).
FT      MOD_RES 39 39 HYDROXYLATION (BY SIMILARITY).
FT      MOD_RES 51 51 HYDROXYLATION (BY SIMILARITY).
FT      MOD_RES 54 54 HYDROXYLATION (BY SIMILARITY).
FT      MOD_RES 57 57 HYDROXYLATION (BY SIMILARITY).
FT      MOD_RES 60 60 HYDROXYLATION (BY SIMILARITY).
FT      MOD_RES 63 63 HYDROXYLATION (BY SIMILARITY).
FT      MOD_RES 75 75 HYDROXYLATION (BY SIMILARITY).
FT      MOD_RES 81 81 HYDROXYLATION (BY SIMILARITY).
FT      MOD_RES 84 84 HYDROXYLATION (BY SIMILARITY).
FT      MOD_RES 90 90 HYDROXYLATION (BY SIMILARITY).
FT      MOD_RES 96 96 HYDROXYLATION (BY SIMILARITY).
FT      MOD_RES 99 99 HYDROXYLATION (BY SIMILARITY).
SQ      SEQUENCE 253 AA; 26589 MW; 1CB40622571BFC9B CRC64;

Query Match. 10.2%; Score 176.5; DB 1; Length 253;
Best Local Similarity 28.2%; Pred. No. 3e-07;
Matches 50; Conservative 40; Mismatches 66; Indels 21; Gaps 8;

QY      2 LPLLGLLGPACWALGP--TPGP-----GSSELR-----AFSAARTTPLEGTSEMA 47
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB      80 IPGIPGKVGPKG--PVGPKGAGPGRPGRPKGSGDYKATQKVAFAFSAALRTVNSALRPNQ 137
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

QY      48 VTFDKVYVNIIGDDEVDVATGQFRCRVPGAYFFSFTAGKAPHKSLSVMLVRNRDEVQ-ALAF 106
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB      138 IRFEKVTITVNDVNDYEPDRSGKFTCKVPGLYFTYHASSRGNLCVNIIVRGDRDRMQKVLTF 197
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

QY      107 DEQRRPGARRAASQASAMLQLDYDGTVWLRLHGAPHYALGAPGAT--FSGYLVYADAD 161
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB      198 CDYAQ-NTFQVTTGGVVLKLEQEEVHLQATD-KNSLLGVEGANSIFTGFLLPMD 252
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

RESULT 15
COLF LEPMA STANDARD; PRT; 419 AA.
AC P98085; Q91080;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Inner ear-specific collagen precursor (Saccular collagen).
OS Lepomis macrochirus (Bluegill).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

```


OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Centrarchidae; Lepomis.
OX NCBI_TaxID=13106;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=95167486; Pubmed=7863331;
RA Davis J.G., Oberholzer J.C., Burns F.R., Greene M.I.;
RT "Molecular cloning and characterization of an inner ear-specific
RT structural protein.";
RL Science 267:1031-1034 (1995).
RN (2)
RP CONCEPTUAL TRANSLATION.
RA Gibson T.;
RL Submitted (MAR-1995) to the SWISS-PROT data bank.
CC -1- FUNCTION: FORMS A MICROSTRUCTURAL MATRIX WITHIN THE OTOLITHIC
CC MEMBRANE (PROBABLY).
CC -1- TISSUE SPECIFICITY: SPECIALIZED SECRETORY SUPPORTING CELLS AT THE
CC OUTER PERIMETER OF THE SACCULAR EPITHELIUM.
CC -1- SIMILARITY: Contains 1 C1Q domain.
CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A FRAMESHIFT HAD TO BE
CC INTRODUCED FOR POSITIONS 391-419 SO AS TO MAXIMIZE THE SIMILARITY
CC WITH OTHER SHORT-CHAIN COLLAGENS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U17431; AAA69978.1; ALT_FRAME.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 3.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR ProDom; PD000007; C1q_helix; 2.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Extracellular matrix; Repeat; Collagen; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 419 INNER EAR-SPECIFIC COLLAGEN.
FT DOMAIN 20 57 NONHELICAL REGION (NC2).
FT DOMAIN 58 274 TRIPLE-HELICAL REGION (COL1).
FT DOMAIN 275 419 NONHELICAL REGION (NC1).
FT DOMAIN 272 419 C1Q.
FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 419 AA; 43634 MW; 570CDB9675FC0F39 CRC64;

Query Match 10.1%; Score 173.5; DB 1; Length 419;
Best Local Similarity 32.7%; Pred. No. 9.4e-07;
Matches 56; Conservative 21; Mismatches 59; Indels 35; Gaps 9;

QY 5 LIGLGPAACWALGPTPGSS--ELRSAPSA---ARTTPLEGTSEMAVTFDKVYVNI 58
DB 258 LKGVGRGPR-----GPKGPPGESVEQIRSAFSVGLFPSRSFP--PSPLPVKFDKVFYNGE 309

QY 59 GDFDVATGQFRCRVPGAYFFSF--TAGKAPHSKSL---SVMLVRNRDEVQALAFDEQRRP 112
DB 310 GHWDPTLNKFNVTYPPGVYLFYSYHITVRNRPRAALVNVGVKRLRTRDSLXGQDIDQ---- 365

QY 113 GARRAASQSAMLQLDYDGTWLRL---HGAPHYALGAPGATFSGYLVYAD 159
DB 366 -----ASNLLALHLTDGDQVWLETLRDWNXG--YSSSEDSTFSGFLYPD 409

Search completed: January 12, 2004, 08:12:56
Job time : 13.8204 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:10 ; Search time 45.3114 Seconds
(without alignments)
1873.686 Million cell updates/sec

Title: US-10-085-167-2
Perfect score: 1726
Sequence: 1 MLPLLLGLGPACWALGPT.....LVYPDLAPAPPGIGASELL 329

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_virus:*
 - 16: sp_bacteriaph:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|--------|----------------------|
| 1 | 1718 | 99.5 | 329 | 4 | Q8IV25 | Q8iv25 homo sapien |
| 2 | 1587 | 91.9 | 326 | 11 | Q8R066 | Q8r066 mus musculus |
| 3 | 1003.5 | 58.1 | 205 | 11 | Q9D0W2 | Q9d0w2 mus musculus |
| 4 | 997.5 | 57.8 | 205 | 11 | Q9DCB6 | Q9dcbb6 mus musculus |
| 5 | 230 | 13.3 | 158 | 4 | Q9H667 | Q9h667 homo sapien |
| 6 | 230 | 13.3 | 1077 | 4 | Q8TE71 | Q8te71 homo sapien |
| 7 | 226 | 13.1 | 158 | 11 | Q8KI10 | Q8ki10 mus musculus |
| 8 | 220.5 | 12.8 | 182 | 11 | Q8RIP2 | Q8rip2 mus musculus |
| 9 | 220.5 | 12.8 | 281 | 11 | Q9QXP7 | Q9qxp7 mus musculus |
| 10 | 213 | 12.3 | 243 | 4 | Q8N6P2 | Q8n6p2 homo sapien |
| 11 | 207 | 12.0 | 294 | 11 | Q9D8U4 | Q9d8u4 mus musculus |
| 12 | 205 | 11.9 | 243 | 11 | Q8R002 | Q8r002 mus musculus |
| 13 | 202 | 11.7 | 243 | 11 | Q8K479 | Q8k479 mus musculus |
| 14 | 199 | 11.5 | 243 | 6 | Q95JD7 | Q95jd7 macaca mula |
| 15 | 196 | 11.4 | 247 | 11 | Q8BRW2 | Q8brw2 mus musculus |
| 16 | 189 | 11.0 | 240 | 6 | Q95MQ4 | Q95mq4 bos taurus |

| | | | | | |
|----|-------|------|------|-----------|--------------------|
| 17 | 188 | 10.9 | 244 | 11 Q8K3R4 | Q8K3R4 rattus norv |
| 18 | 186 | 10.8 | 287 | 11 Q8CFR0 | Q8CFR0 mus musculu |
| 19 | 185.5 | 10.7 | 312 | 11 Q8CHX9 | Q8CHX9 mus musculu |
| 20 | 180.5 | 10.5 | 264 | 11 Q8BKRO | Q8BKRO mus musculu |
| 21 | 179.5 | 10.4 | 120 | 11 Q8R1Z2 | Q8R1Z2 mus musculu |
| 22 | 176.5 | 10.2 | 289 | 11 Q8BVD7 | Q8BVD7 mus musculu |
| 23 | 170 | 9.8 | 194 | 6 Q95J95 | Q95J95 canis fami1 |
| 24 | 168 | 9.7 | 245 | 11 Q9DCM6 | Q9DCM6 mus musculu |
| 25 | 167.5 | 9.7 | 744 | 11 Q92IS8 | Q92IS8 mus musculu |
| 26 | 167.5 | 9.7 | 744 | 11 Q8BGL6 | Q8BGL6 mus musculu |
| 27 | 166 | 9.6 | 675 | 6 Q9N178 | Q9N178 sus scrofa |
| 28 | 164 | 9.5 | 295 | 11 Q9Z1K4 | Q9Z1K4 rattus norv |
| 29 | 163 | 9.4 | 224 | 4 Q8ITK8 | Q8ITK8 homo sapien |
| 30 | 163 | 9.4 | 224 | 11 Q8BGU2 | Q8BGU2 mus musculu |
| 31 | 162.5 | 9.4 | 246 | 11 Q9ES30 | Q9ES30 mus musculu |
| 32 | 161 | 9.3 | 173 | 6 Q62789 | Q62789 sus scrofa |
| 33 | 155 | 9.0 | 196 | 11 Q920N0 | Q920N0 tamias sibi |
| 34 | 152 | 8.8 | 197 | 11 Q9JHG0 | Q9JHG0 mus musculu |
| 35 | 150 | 8.7 | 333 | 4 Q8ITU4 | Q8ITU4 homo sapien |
| 36 | 145.5 | 8.4 | 705 | 4 Q8TEJ5 | Q8TEJ5 homo sapien |
| 37 | 141.5 | 8.2 | 1017 | 11 Q99KA1 | Q99KA1 mus musculu |
| 38 | 140 | 8.1 | 198 | 11 Q8BMF0 | Q8BMF0 mus musculu |
| 39 | 139 | 8.1 | 198 | 11 Q8BME9 | Q8BME9 mus musculu |
| 40 | 137.5 | 8.0 | 195 | 11 Q8BZS3 | Q8BZS3 mus musculu |
| 41 | 134.5 | 7.8 | 213 | 5 P83425 | P83425 mytilus edu |
| 42 | 132.5 | 7.7 | 347 | 4 Q96IH6 | Q96IH6 homo sapien |
| 43 | 131.5 | 7.6 | 185 | 5 Q9GQV4 | Q9GQV4 strongyloce |
| 44 | 131.5 | 7.6 | 583 | 4 Q96G58 | Q96G58 homo sapien |
| 45 | 131.5 | 7.6 | 992 | 4 Q9UG76 | Q9UG76 homo sapien |

ALIGNMENTS

RESULT 1

Q8IV25

ID Q8IV25 PRELIMINARY; PRT; 329 AA.

AC Q8IV25;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Similar to C1q and tumor necrosis factor related protein 4.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA Strausberg R.;

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC035628; AAH35628.1; -

SQ SEQUENCE 329 AA; 35256 MW; 16064DA8182A6732 CRC64;

| | | | | |
|-----------------------|--------------|---|---------------|-------------|
| Query Match | 99.5%; | Score 1718; | DB 4; | Length 329; |
| Best Local Similarity | 99.7%; | Pred. No. 2.2e-134; | | |
| Matches 328; | Conservative | 0; | Mismatches 1; | Indels 0; |
| Gaps | 0; | | | |
| QY | 1 | MLPLLLGLGPACWALGPTPGSSSELSRASFSAARTTPLEGTSEMAVTFDKVYVNI | 60 | |
| DB | 1 | MLPLLLGLGPACWALGPTPGSSSELSRASFSAARTTPLEGTSEMAVTFDKVYVNI | 60 | |
| QY | 61 | FDVATGFCRCRVPGAYFFSFAGKAPHSLSVMLVRNRDEVQALAFDEQRRPGARRASQ | 120 | |
| DB | 61 | FDVATGFCRCRVPGAYFFSFAGKAPHSLSVMLVRNRDEVQALAFDEQRRPGARRASQ | 120 | |
| QY | 121 | SAMQLDYGDTVMRLHGAPHYALGAPATFSGYLVYADADADAPARGPPAPPEPRSAFS | 180 | |
| DB | 121 | SAMQLDYGDTVMRLHGAPHYALGAPATFSGYLVYADADADAPARGPPAPPEPRSAFS | 180 | |
| QY | 181 | AARTSLVGSADAGPGRHQPPLAFDTEFVNIGGDFDAAGVFRCLPGAYFFSFTLGKLP | 240 | |
| DB | 181 | AARTSLVGSADAGPGRHQPPLAFDTEFVNIGGDFDAAGVFRCLPGAYFFSFTLGKLP | 240 | |

OY 241 KTLVKLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRRGDVWLLSHDHGCGAYSNH 300
 DB 241 KTLVKLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRRGDVWLLSHDHGCGAYSNH 300

OY 301 GKTYTSGFLVYPDLAPAPPGLGASELL 329
 DB 301 GKTYTSGFLVYPDLAPAPPGLGASELL 329

RESULT 2

O8R066 PRELIMINARY; PRT; 326 AA.

DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Similar to C1q and tumor necrosis factor related protein 4.

GN 0710001E10RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Strauberg R.;

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC027315; AAH27315.1; -

DR MGD; MGI:1914695; 0710001E10RIK.

DR InterPro; IPR001073; C1q.

DR Pfam; PF00386; C1q; 2.

DR SMART; SM00110; C1Q; 2.

DR PROSITE; PS01113; C1Q; 2.

SQ SEQUENCE 326 AA; 35057 MW; 72339172B7B1051A CRC64;

Query Match 91.9%; Score 1587; DB 11; Length 326;
 Best Local Similarity 94.2%; Pred. No. 1.6e-123;
 Matches 309; Conservative 2; Mismatches 15; Indels 2; Gaps 2;

OY 1 MLPLILGLGPACWALGPTPGSSSELSAFAAARTTPLEGTSEMAVTFDKVVYVNIQD 60

DB 1 MLLLLIGFLGPACWALGPA-GPGSSSELSAFAAARTTPLEGTSEMAVTFDKVVYVNIQD 59

OY 61 FDVATGQFRCRVPGAYFFSFTAGAPHSLSVMLVRNRDEVQALAFDEQRRPGARRASQ 120

DB 60 FDAATGFRRCRVPGAYFFSFTAGAPHSLSVMLVRNRDEVQALAFDEQRRPGARRASQ 119

OY 121 SAMLOLDYDVTWMLRHGAPHYALGAPGATFSGYLVYADADADAPARGPPAPPEPRSAFS 180

DB 120 SAMLOLDYDVTWMLRHGAPHYALGAPGATFSGYLVYADADADAPARG-PAAPEPRSAFS 178

OY 181 AARTSLVGSADAGPGRHQPPLAFDTEFVNIGDGFDAAGVFRCLPGAYFFSFTLGKLP 240

DB 179 AARTSLVGSADAGPGRHQPPLAFDTEFVNIGDGFDAAGVFRCLPGAYFFSFTLGKLP 238

OY 241 KTLVKLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRRGDVWLLSHDHGCGAYSNH 300

DB 239 KTLVKLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRRGDVWLLSHDHGCGAYSNH 298

OY 301 GKTYTSGFLVYPDLAPAPPGLGASEL 328

DB 299 GKTYTSGFLVYPDLAPAPPGLGASEL 326

RESULT 3

O9D0W2 PRELIMINARY; PRT; 205 AA.

AC Q9D0W2; 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

DE 0710001E10RIK protein.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Embryo;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,

RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Guetlich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyo-oka K., Wang K.H., Weltz C., Whittaker C., Wilming L.,

RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,

RA Hayashizaki Y.;

RA "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

DR EMBL; AK004340; BAB23268.1; -

DR MGD; MGI:1914695; 0710001E10RIK.

DR InterPro; IPR001073; C1q.

DR Pfam; PF00386; C1q; 1.

DR PRINTS; PR00007; COMPLEMENTC1Q.

DR SMART; SM00110; C1Q; 1.

DR PROSITE; PS01113; C1Q; 1.

SQ SEQUENCE 205 AA; 22190 MW; B9B237793C54786D CRC64;

Query Match 58.1%; Score 1003.5; DB 11; Length 205;
 Best Local Similarity 93.7%; Pred. No. 2e-75;
 Matches 193; Conservative 1; Mismatches 11; Indels 1; Gaps 1;

OY 123 MLQLDYDVTWMLRHGAPHYALGAPGATFSGYLVYADADADAPARGPPAPPEPRSAFA 182

DB 1 MLQLDYDVTWMLRHGAPHYALGAPGATFSGYLVYADADADAPARG-PAAPEPRSAFA 59

OY 183 RTRSLVGSADAGPGRHQPPLAFDTEFVNIGDGFDAAGVFRCLPGAYFFSFTLGKLP 242

DB 60 RTRSLVGSADAGPGRHQPPLAFDTEFVNIGDGFDAAGVFRCLPGAYFFSFTLGKLP 119

OY 243 LSVKLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRRGDVWLLSHDHGCGAYSNH 302

DB 120 LSVKLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRRGDVWLLSHDHGCGAYSNH 179

OY 303 YTFSGFLVYPDLAPAPPGLGASEL 328

DB 180 YTFSGFLVYPDLAPAPPGLGASEL 205

RESULT 4

O9DCB6 PRELIMINARY; PRT; 205 AA.

AC Q9DCB6; 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE 0710001E10RIK protein.

GN 0710001E10RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boftelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzairelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK002948; BAB22473.1; -
DR MGD; MGI:1914695; 0710001E10Rik.
DR InterPro; IPR001073; Clq.
DR Pfam; PF00386; Clq; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
SQ SEQUENCE 205 AA; 22215 MW; 56AD37793C437300 CRC64;

Query Match 57.8%; Score 997.5; DB 11; Length 205;
Best Local Similarity 93.2%; Pred. No. 6.4e-75;
Matches 192; Conservative 1; Mismatches 12; Indels 1; Gaps 1;
QY 123 MLQLDYGDYTWLRLHGAHYALGAFGATFSGYLVYADADADAPARGPAPPEPRSAFSA 182
DB 1 MLQLDYGDYTWLRLHGAHYALGAFGATFSGYLVYADADADAPARGPAPPEPRSAFSA 59
QY 183 RTRSLVSDAGPGRHQPPLAFDTEFVNIGDFAAGVFRCLPGAYFFSFTLGKLPKRT 242
DB 60 RTRSLVSDAGPGRHQPPLAFDTEFVNIGDFAAGVFRCLPGAYFFSFTLGKLPKRT 119
QY 243 LSVKLMKRNDEVQAMITYDGDASRRREMOSQSVMLALRGDAVWLSHDHGXYGAYSNHGK 302
DB 120 LSVKLMKRNDEVQAMITYDGDASRRREMOSQSVMLALRGDAVWLSHDHGXYGAYSNHGK 179
QY 303 YITFSGFLVYPDLPAPAPGLGASEL 328
DB 180 YITFSGFLVYPDLPAPAPGLGASEL 205

RESULT 5
Q9H667 PRELIMINARY; PRT; 158 AA.
AC Q9H667;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Hypothetical protein FLJ22569.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isegai T., Sugano S.,
RT "NEDO human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]

RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK026222; BAB15398.1; -
DR EMBL; BC007520; AAH07520.1; -
DR InterPro; IPR001073; Clq.
DR Pfam; PF00386; Clq; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Hypothetical protein.
SQ SEQUENCE 158 AA; 17625 MW; 47DB10EDD6DC9760 CRC64;
Query Match 13.3%; Score 230; DB 4; Length 158;
Best Local Similarity 38.4%; Pred. No. 2e-11;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;
QY 163 DAPARGP-----PAPPEPRSAFSAARTSLVSDAGPGRHQPPLAFDTEFVNIG 212
DB 5 DVEPTNPAATILPVHVPLPQGMRAVFAFSAART-----SNLAPGTLDDPIVFDLLNNGE 59
QY 213 DFDAAAGVFRCLPGAYFFSFTLGKLP-RKTLVSKLMKRNDEVQAMITYDGDASRRREMOS 271
DB 60 TFDLQGRFNCVPNGTYVFIFHMLKLA VNPVLYNLMKNEEVLVSAYANDGAP-DHETAS 118
QY 272 QSVMLALRRGDAVWLSHDHGXYGA-YSNHGKYITFSGFLVYPD 314
DB 119 NHAIIQLFGDQIWLRLH---RGAITYGSSWKYSTFSGYLLYQD 158

RESULT 6
Q8TE71 PRELIMINARY; PRT; 1077 AA.
ID Q8TE71;
AC Q8TE71;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE EGG1L.
GN EGG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Aerbaajina W., Miller J.L.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY074490; AAL71549.1; -
DR InterPro; IPR001073; Clq.
DR Pfam; PF00386; Clq; 1.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
SQ SEQUENCE 1077 AA; 120974 MW; 2B88BF3C47D032D6 CRC64;

Query Match 13.3%; Score 230; DB 4; Length 1077;
Best Local Similarity 38.4%; Pred. No. 2.4e-10;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;
QY 163 DAPARGP-----PAPPEPRSAFSAARTSLVSDAGPGRHQPPLAFDTEFVNIG 212
DB 924 DVEPTNPAATILPVHVPLPQGMRAVFAFSAART-----SNLAPGTLDDPIVFDLLNNGE 978
QY 213 DFDAAAGVFRCLPGAYFFSFTLGKLP-RKTLVSKLMKRNDEVQAMITYDGDASRRREMOS 271
DB 979 TFDLQGRFNCVPNGTYVFIFHMLKLA VNPVLYNLMKNEEVLVSAYANDGAP-DHETAS 1037
QY 272 QSVMLALRRGDAVWLSHDHGXYGA-YSNHGKYITFSGFLVYPD 314
DB 1038 NHAIIQLFGDQIWLRLH---RGAITYGSSWKYSTFSGYLLYQD 1077

RESULT 7


```

Q8K110 ID Q8K110 PRELIMINARY; PRT; 158 AA.
AC Q8K110;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Similar to hypothetical protein FLJ22569.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027523; AAH27523.1; -.
DR InterPro; IPR001073; Clq.
DR Pfam; PF00386; Clq; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KM Hypothetical protein.
SQ SEQUENCE 158 AA; 17533 MW; 86E9321C99225FCB CRC64;

```

| | | | | |
|-----------------------|--------|--------------------|----------------|-------------|
| Query Match | 13.1%; | Score 226; | DB 11; | Length 158; |
| Best Local Similarity | 39.1%; | Pred. No. 4.2e-11; | | |
| Matches | 63; | Conservative 18; | Mismatches 68; | Indels 12; |
| | | | | Gaps 5; |

QY 156 VYADADADAEARGPPAPPEPRSAFSAARTSLVGS DAGPGRHQPLAFDTEFVNIGDFD 215
| : | | : | | | | : | | :
DB 8 VTSPAAILPEVHIYPLPQQMRVAFSAART-----SNLAPGTLDPQIVFDLLNNIGETFN 62

QY 216 AAGVFRCLLEGAYFFSFTLGKLP-RKTLSVKLMKNRDEVOAMIYDDGASRRREQSOVS 274
| | : | | : | | : : :
Db 63 LQLGRFNCPVNGTYVFIFHMLKLAVNPLYYNLMKNEEVLVSAYANDGAR-DHETASNHA 121

QY 275 MLALRRGDVWLLSHDHGCGYA-YSNHGKYYITPGEFLVPD 314
: | : | : | : | : | : | : | : | : | : | : | : | : |
Db 122 VLQLLGGDIWLRLH---RGAIYGSSWKYSTFSGVLRYQD 158

| | |
|-----------|---------------------------|
| RESULT 8 | |
| Q8R1P2 | |
| ID Q8R1P2 | PRELIMINARY; PRT; 182 AA. |

DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Similar to RIKEN CDNA 1600017K21 gene (Fragment).
GN 1600017K21RIK.

OS Mus musculus (Mouse) .
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus
 OX NCBI_TaxID=10090;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Strausberg R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases

DR EMBL; BC023468; AAH23468.1; -
DR MGD; MGI:1919254; 1600017K2IRik.
DR InterPro; IPR001073; Clq.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; Clq; 1.
DR SMART; SM00110; Clq; 1.
FT NON TER 1 1
SQ SEQUENCE 182 AA; 20863 MW; 559C73DDE9517882F CRC64;

| | | | | |
|-----------------------|------------------|--------------------|-----------|-------------|
| Query Match... | 12.8%; | Score 220.5; | DB 11; | Length 182; |
| Best Local Similarity | 39.2%; | Pred. No. 1.4e-10; | | |
| Matches 56; | Conservative 20; | Mismatches 58; | Indels 9; | Gaps 4; |

QY 177 SAESAARTRSLVGS DAGPGPRHQPLAEDTEFVNIGGDFDAAAGVFRCLRLPGA YFFSFTLG 236
 : ||| | :: | : | : ||||| : | : | : ||| :
 Db 47 AAFSVGRKKALHSND-----YEQPVVFDETFEVLNLYKHFMFTGKFYCYVPGLYFFSLNVH 101

QY 237 KLPRKTLVKLMKNRDEVQA MIYDDGASRRREMOSQSVM LALRGDA VW--LISHDHGY 294
 : | : : ||| : || : | ||||| : | : | : : :
 Db 102 TWMNQKETYLHI MKNEBEV-VILYAQ-VSDRSIMQSQSI MMELREDEDE V WVRLEFKGERENA 159

QY 295 GAYSNHGKYITFSGFLVYPDLAP 317
 ||||| : ||| |
 Db 160 IFSDEFDTYITFSGYLVKPASEP 182

```

RESULT 9
O9QXP7
ID O9QXP7 PRELIMINARY; PRT; 281 AA.
AC O9QXP7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative secreted protein ZSIG37 (1600017K21RIK protein).
GN ZSIG37 OR 1600017K21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Shepard P., Deisher T., Grant F., Chen L., Haldeman B., McKnight G.,
RA Whitmore T., O'Hara P.;
RT "Mus musculus putative secreted protein.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Placenta;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamlya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AF192499; AAF06664.1; -.
DR EMBL; AK005484; BAB24070.1; -.
DR MGD; MGI:1919254; 1600017K21Rik.
DR InterPro; IPR001073; Clq.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; Clq; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
SQ SEQUENCE 281 AA; 32009 MW; C9816216DB6419E2 CRC64;

```

```

Query Match          12.8%; Score 220.5; DB 11; Length 281;
Best Local Similarity 39.2%; Pred. NO. 2.5e-10;
Matches 56; Conservative 20; Mismatches 58; Indels 9; Gaps 4

QY 177 SAFSARTSLVGS DAGPGRHQPLAFDTEFVNIGD FDDAAGVRCRLPGAYFFSFTLG 236
:||||| :::: : |||: ||||| : : : ||| ||||| :

```

Db 146 AAFSVGRKKALHSND-----YEQPVFDTFVNLYKHFNMFTGKFCYVPGIYFFSLNVH 200
QY 237 KLPRKTLVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRGDAVM--LLSHDHDGY 294
Db 201 TWNQKETYLHIMKNEEV-VILYAQ-VSDRSIMOSQSLMELREDEVMVRLFRGERENA 258
QY 295 GAYSNHGKIYTFSGFLVYPDLAP 317
Db 259 IFSDEFDTYITFSGYLVKPASEP 281

RESULT 10

Q8N6P2 PRELIMINARY; PRT; 243 AA.
AC Q8N6P2;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE C1q and tumor necrosis factor related protein 5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC029485; AAH29485.1; -
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
SQ SEQUENCE 243 AA; 25326 MW; 6D9306A0EB21B44A CRC64;

Query Match 12.3%; Score 213; DB 4; Length 243;
Best Local Similarity 34.9%; Pred. No. 8.8e-10;
Matches 66; Conservative 18; Mismatches 83; Indels 22; Gaps 6;

QY 138 GAPHYALGAPGATFSG---YLIVYADAD---ADAPARGPAPP-----EPRSAPSAARTR 185
Db 57 GAP---GAPGEKGEGRPGLPGPRGDPGRGEAGPAGPTGPAGECSVPRSAFSAKRSE 112
QY 186 SLVGS DAGPGRHQPLAFDTEFVNIGDFDAAGVFRCLPGAYFFSFTLGLPRKTLSTV 245
Db 113 SRV-----PPSDAPLPFDRVLVNEQGHYDAVTGKTCQVPGVYFA-VHATVYRASLQF 166
QY 246 KLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRGDAVMVLLSHDHDGYGAYSNHGKIY 305
Db 167 DLVKNGESIASFPQFGGWPKPASLSGGAMVRLPEPDQVWVQVGVDYIGIYASIKTDS 226
QY 306 FSGFLVYPD 314
Db 227 FSGFLVYSD 235

RESULT 11

Q9D8U4 PRELIMINARY; PRT; 294 AA.
AC Q9D8U4;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE 1810033K05Rik protein (RIKEN CDNA 1810033K05 gene).
GN 1810033K05RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido T., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Wetz C., Wittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK007683; BAB25187.1; -
DR EMBL: BC030324; AAH30324.1; -
DR MGD; MGI:1916433; 1810033K05Rik.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 2.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR PROSITE; PS01113; C1Q; 1.
SQ SEQUENCE 294 AA; 30865 MW; 6D3905AE7C19E6FA CRC64;

Query Match 12.0%; Score 207; DB 11; Length 294;
Best Local Similarity 33.7%; Pred. No. 3.6e-09;
Matches 55; Conservative 31; Mismatches 47; Indels 30; Gaps 8;

QY 7 GLLGPAACWALGPTPGGSSELRSAFSA--RTPLEGTSEMAVTFDKVYNIGDFDVA 64
Db 145 GLPGPCSC-----GSSRAKSAFSAVATKSYPRE--RLPIKFDKILMNEGHYNAS 192
QY 65 TGQFRCRVPVPAFFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPARRAASQASML 124
Db 193 SGKFCVSPGIIYFTYDITLA-NKHLATGLVHN-GQYRIRTFD-ANTGNHDAVSGSTIL 248
QY 125 QLDYGDYVRLHGA-----PHYALGAPATFSGYLVYAD 159
Db 249 ALKEGDEVWLQIFYSEQNGLFYDPYWT---DSLFTGFLIYAD 287

RESULT 12

Q8R002 PRELIMINARY; PRT; 243 AA.
AC Q8R002;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Similar to DKFZPS68B0621 protein (Hypothetical 25.4 kDa protein).
GN C1QTNF5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

RA Strausberg R.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases
 DR EMBL; BC023068; AAH23068.1; -.
 DR EMBL; BC025174; AAH25174.1; -.
 DR MGD; MGI:2385958; C1qcnf5.
 DR InterPro; IPR001073; C1q.
 DR InterPro; IPR000087; Collagen.
 DR Pfam; PF00386; C1q; 1.
 DR Pfam; PF01391; Collagen; 1.
 DR SMART; SM00110; C1Q; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 243 AA; 25420 MW; 498129CD051DB97B CRC64;

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 11.9% | Score 205; | DB 11; | Length 243; |
| Best Local Similarity | 32.8%; | Pred. No. 4.1e-09; | | |
| Matches 64; | Conservative 16; | Mismatches 81; | Indels 34; | Gaps 6; |

```
QY      138 GAPHYALGAPGATFSGYLVTADADADAPA-RGPAP-----PEPRSAF 179
          |||   ||||   |               |||||
Db      57 GAP----GAPEGKEG-----GRGLDPGRGEPPRGEGAPMGAI GPAGECSVPPRSAF 106
          |||   ||||   |               |||||

QY      180 SAARTSLVGSDAGPGPRHQPLAFDTEFNIGDFDAAGVFCRCRLPGAYFFSFTLGKLP 239
          ||| : | | | | | | | : | | | | | | | : | | | | | | | :
Db      107 SAKRSES RV-----PPPADTFLPFDRVLN EQGHFDPTTGKFTCCQVPGVYYFA-VHATVY 160
          ||| : | | | | | | | : | | | | | | | : | | | | | | | :

QY      240 RKTLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRRGDAVWLLSHDHGYGAYS N 299
          | : | | | | | : | : | | | | | : | | | | | | | : | | :
Db      161 RASLQFDLVKNGQSIA SFQYFGGWPKPASLSGAMVRLEPEDQVWVQVGVDYIGIYAS 220
          | : | | | | | : | : | | | | | : | | | | | | | : | | :

QY      300 HGKYITFSGF LVYPD 314
          ||||| ||||| |
Db      221 IKTDSTFSGF LVYSD 235
```

| RESULT 13 | | | |
|-----------|--|---|--------------|
| ID | Q8K479 | PRELIMINARY; | PRT; 243 AA. |
| AC | Q8K479; | | |
| DT | 01-OCT-2002 | (TREMBLrel. 22, Created) | |
| DT | 01-OCT-2002 | (TREMBLrel. 22, last sequence update) | |
| DT | 01-MAR-2003 | (TREMBLrel. 23, last annotation update) | |
| DE | Complement-c1q tumor necrosis factor-related protein. | | |
| GN | C1QTNF5. | | |
| OS | Mus musculus (Mouse). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | |
| OX | NCBI_TaxID=10090; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=CS7BL/6J; | | |
| RX | MEDLINE=22135657; Pubmed=12140190; | | |
| RA | Kameya S., Hawes N.L., Chang B., Heckenlively J.R., Naggert J.K., | | |
| RA | Nishina P.M.; | | |
| RT | "Mfrp, a gene encoding a frizzled related protein, is mutated in the | | |
| RT | mouse retinal degeneration 6."; | | |
| RL | Hum. Mol. Genet. 11:1879-1886(2002). | | |
| DR | EMBL; AF469650; AAM89217.1; - | | |
| DR | MGD; MGI:2385958; C1qtnf5. | | |
| DR | InterPro; IPR001073; C1q. | | |
| DR | InterPro; IPR000087; Collagen. | | |
| DR | Pfam; PF00386; C1q; 1. | | |
| DR | Pfam; PF01391; Collagen; 2. | | |
| DR | PRINTS; PR00007; COMPLEMENTC1Q. | | |
| DR | SMART; SM00110; C1Q; 1. | | |
| SEQ | SEQUENCE 243 AA; 25436 MW; 9F4D5804349791D9 CRC64; | | |

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 11.7%; | Score 202; | DB 11; | Length 243; |
| Best Local Similarity | 32.3%; | Pred. No. 7.2e-09; | | |
| Matches 63; | Conservative 17; | Mismatches 81; | Indels 34; | Gaps 6; |

```

QY      138 GAPHYALGAPGATFSGYLVYADADADAPA-RGPAP-----EPPRSAF 179
          |||      |||      |      |||      |||
Db      57 GAP---GAPDEKGGC-----GRPGLPGPRGGCPRGCEAGPMGAIGPAGECCSVPPRSAF 106

```

QY 180 SAARTSLVSGDAGPGPRHQPLAEDTEFVNIGGDFDAAGVFRCLPGAYFFSFTLGKLP 239
 107 SAKRSESrv-----PPPADTPLPFDRLVLINEQGHYDPTTGKFTCQVPGYFFA-VHATVY 160

QY 240 RKTLSVKLMKNRDEVQAMIYDDGASRRREMOSQSVMALRLRGDAVWLSDHDGYGAYSN 299
 161 RASLQFDLVKNGQSIAFFQYFGGWPKPASLSGAMVRLPEPDQVWVGVGVDYIGIYAS 220

QY 300 HGKYITFSGFLVYPD 314
 221 IKTDSTFSGFLVYSD 235

Db

| | |
|-----------|--------------------------|
| RESULT 14 | |
| Q95JD7 | |
| ID Q95JD7 | PRELIMINARY; PRT; 243 AA |

DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Adiponectin.
GN APM1.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adipose tissue;
RX MEDLINE=21232234; PubMed=11334417;
RA Hotta K., Funahashi T., Bodkin N.L., Ortmeier H.K., Arita Y.,
RA Hansen B.C., Matsuzawa Y.;
RT "Circulating concentrations of the adipocyte protein adiponectin are
RT decreased in parallel with reduced insulin sensitivity during the
RL progression to type 2 diabetes in rhesus monkeys.";
RL Diabetes 50:1126-1133(2001).
DR EMBL; AF404407; AAK92202.1; -.
DR InterPro; IPR001073; Clq.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; Clq; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PRO0007; COMPLEMENTC1Q.
DR PRODom; PD000007; Collagen; 1.
DR SMART; SM00110; Clq; 1.
DR PROSITE; PS01113; Clq; 1.
DR Collagen.
KW SEQUENCE.
SQ 243 AA; 26264 MW; 49A45DAF2B4613FD CRC64;

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 11.5%; | Score 199; | DB 6; | Length 243; |
| Best Local Similarity | 29.7%; | Pred. No. 1.3e-08; | | |
| Matches 54; | Conservative 31; | Mismatches 61; | Indels 36; | Gaps 8 |

```

QY      7 GLIGP-----AACWALGP-----TPGSSSELRSAFSAARTTPLEGTSEMAV 48
        ||::|||          |||          |||          |||          |
Db      71 GLIGPKDGTGETGVGTGAEGPRGPGIQRKKEPEGEAYVYR$AFSGVLETVT-VPNMPI 129
        ||::|||          |||          |||          |||          |
QY      49 TFDKYVNIGGEDVDVATGOFRCRYGAYFFSFTAGAKPH----KSLSVMLVRNRDEVQA 103
        ||::|||          :||:|||||:||:|||||:||:|||||
Db      130 RFTKIFYNQNHYYDGSTGKFHCNIPGLYFAY-----HITVYMKVSVL$FK-KDKAML 182
        ||::|||          :||:|||||:||:|||||:||:|||||
QY      104 LAFDEQRRPGARRAASQSAMLQLDPYGDVTWLRLHGAPH---YALGABGATEFSGYLVYAD 159
        :||:|||||          :||:|||||:||:|||||:||:|||||
Db      183 FTYDYQENNVDD-Q-ASGSVLHLHLEVDAQVWLQVYGEGERNGLYADNDNDSTFTGFLLYHD 241
        :||:|||||          :||:|||||:||:|||||:||:|||||
QY      160 AD 161
        :
Db      242 TN 243

```

RESULT 15
Q8BRW2

```

ID      Q8BRW2      PRELIMINARY;      PRT;      247 AA.
AC      Q8BRW2;
DT      01-MAR-2003 (TREMBLrel. 23, Created)
DT      01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT      01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE      Adipocyte complement related protein of 30 kDa.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX      MEDLINE=22354683; PubMed=12466851;
RA      The FANTOM Consortium,
RA      the RIKEN Genome Exploration Research Group Phase I & II Team;
RT      "Analysis of the mouse transcriptome based on functional annotation of
RT      60,770 full-length cDNAs.";
RL      Nature 420:563-573 (2002).
DR      EMBL; AK041214; BAC30866.1; -.
SQ      SEQUENCE 247 AA; 26751 MW; 0D3FA64C789CAEF3 CRC64;

```

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 11.4%; | Score 196; | DB 11; | Length 247; |
| Best Local Similarity | 31.9%; | Pred. No. 2.3e-08; | | |
| Matches 58; | Conservative 30; | Mismatches 58; | Indels 36; | Gaps 11; |

| | | | |
|----|-----|--|-----|
| QY | 7 | GILGP-----AACWALGP-----TPG-----PGSSE--LRSAFSAARTTPLEGTSEMAY | 48 |
| Db | 75 | GLIGPKGETGVGMTGAEGPRGFPTGRKGEPEAAVYRRSAFSGLETRVT-VENVPI | 133 |
| QY | 49 | TFDKYVNIGDFDVATGQFRCRVPAGVFESFTAGKAPH-----KSLSVMLVRNRDEVQA | 103 |
| Db | 134 | RFTKIIFYNQNMHYDGGSTGKFYCNIPLLYFSY-----HITVMKDVKVSLEFK-KDKAVL | 186 |
| QY | 104 | LAFDEQRPGARRAASQSAMLQLDYGDYWLRLHG-APH--YALGAPGATFSGYLVIAD | 159 |
| Db | 187 | FTYDQYQEKNVDQ-ASGSVLHLLEVGDQVWLQVYGDDGNGLYADNVNDSTFTGFLLYHD | 245 |
| QY | 160 | AD 161 | |
| Db | 246 | TN 247 | |

Search completed: January 12, 2004, 08:19:07
Job time : 49.3114 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:09 ; Search time 49.5798 Seconds
(without alignments)
1053.272 Million cell updates/sec

Title: US-10-085-167-2

Perfect score: 1726

Sequence: 1 MLPLLLGLGPACWALGPT.....LVYPDLAPAPPGLGASELL 329

Scoring table: BL0SUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03:.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|----------|--------------------|
| 1 | 1726 | 100.0 | 329 | 22 | AAB61606 | Human ZACRP4. Hom |
| 2 | 1707 | 98.9 | 329 | 23 | ABG79643 | Human novel secret |
| 3 | 1666 | 96.5 | 329 | 22 | AAB61424 | Monkey MANGO 245 p |
| 4 | 1567 | 90.8 | 348 | 22 | AAB61423 | Human MANGO 245 pr |
| 5 | 1540 | 89.2 | 239 | 23 | ABG70385 | Adipocyte compleme |
| 6 | 1491 | 86.4 | 334 | 22 | AAB61466 | Human MANGO 245 ma |
| 7 | 1445.5 | 83.7 | 284 | 23 | ABG70384 | Adipocyte compleme |
| 8 | 882 | 51.1 | 199 | 22 | AAB61488 | Murine MANGO 245 p |
| 9 | 760 | 44.0 | 192 | 22 | AAB61479 | Human MANGO 245 cy |

| | | | | | | |
|----|-------|------|------|----|----------|--------------------|
| 10 | 707.5 | 41.0 | 221 | 23 | ABG70383 | Adipocyte compleme |
| 11 | 683 | 39.6 | 134 | 22 | AAB61473 | Monkey MANGO 245 C |
| 12 | 664 | 38.5 | 127 | 22 | ABB11578 | Human secreted pro |
| 13 | 650 | 37.7 | 125 | 22 | AAB61477 | Human MANGO 245 ex |
| 14 | 646 | 37.4 | 126 | 22 | AAB61469 | Human MANGO 245 Cl |
| 15 | 639 | 37.0 | 126 | 22 | AAB61489 | Murine MANGO 245 C |
| 16 | 622 | 36.0 | 126 | 22 | AAB61472 | Monkey MANGO 245 C |
| 17 | 618 | 35.8 | 123 | 21 | AAB61479 | Human ORFX ORF1953 |
| 18 | 609 | 35.3 | 117 | 22 | AAB61470 | Human MANGO 245 Cl |
| 19 | 594 | 34.4 | 133 | 23 | ABG70386 | Adipocyte compleme |
| 20 | 552 | 32.0 | 130 | 22 | AAB61468 | Mature monkey MANG |
| 21 | 237.5 | 13.8 | 252 | 22 | AAB49590 | Human zacrp5. Hom |
| 22 | 237.5 | 13.8 | 252 | 23 | AAO19412 | Human molecule for |
| 23 | 237.5 | 13.8 | 252 | 23 | AAU83100 | Novel secreted pro |
| 24 | 237.5 | 13.8 | 252 | 23 | ABB53292 | Human polypeptide |
| 25 | 231 | 13.4 | 800 | 22 | AAM78334 | Human protein SEQ |
| 26 | 230 | 13.3 | 202 | 22 | AAM99927 | Human polypeptide |
| 27 | 230 | 13.3 | 709 | 23 | ABB08645 | Human pancreatic C |
| 28 | 230 | 13.3 | 710 | 23 | AAO15420 | Human gensec metab |
| 29 | 230 | 13.3 | 746 | 20 | AAI29512 | Human lung tumour |
| 30 | 230 | 13.3 | 746 | 21 | AAB44461 | Human lung tumour |
| 31 | 230 | 13.3 | 746 | 22 | AAE13802 | Human lung tumour |
| 32 | 230 | 13.3 | 908 | 22 | AAE09839 | Novel human protei |
| 33 | 230 | 13.3 | 909 | 22 | AAE09845 | Novel human protei |
| 34 | 230 | 13.3 | 957 | 22 | AAE09841 | Novel human protei |
| 35 | 230 | 13.3 | 958 | 22 | AAE09847 | Novel human protei |
| 36 | 230 | 13.3 | 992 | 22 | AAE09843 | Novel human protei |
| 37 | 230 | 13.3 | 993 | 22 | AAE09849 | Novel human protei |
| 38 | 230 | 13.3 | 1042 | 22 | AAE09838 | Novel human protei |
| 39 | 230 | 13.3 | 1043 | 22 | AAE09844 | Novel human protei |
| 40 | 230 | 13.3 | 1091 | 22 | AAE09840 | Novel human protei |
| 41 | 230 | 13.3 | 1092 | 22 | AAE09846 | Novel human protei |
| 42 | 230 | 13.3 | 1126 | 22 | AAE09842 | Novel human protei |
| 43 | 230 | 13.3 | 1127 | 22 | AAE09848 | Novel human protei |
| 44 | 220.5 | 12.8 | 281 | 20 | AAW98014 | Mouse adipocyte-sp |
| 45 | 220.5 | 12.8 | 281 | 21 | AAB08428 | Amino acid sequenc |

ALIGNMENTS

| | | | | | | |
|----------|--|----------------------------|--|--|--|--|
| RESULT 1 | | | | | | |
| ID | AAB61606 | standard; Protein; 329 AA. | | | | |
| XX | XX | | | | | |
| AC | AAB61606; | | | | | |
| XX | XX | | | | | |
| DT | 05-APR-2001 | (first entry) | | | | |
| XX | XX | | | | | |
| DE | Human ZACRP4. | | | | | |
| XX | XX | | | | | |
| KW | Human; zacrp4; complement factor C1q domain; chromosome 11q11; | | | | | |
| KW | energy balance; cellular metabolic reaction; autocrine factor; | | | | | |
| KW | development; cell proliferation; differentiation; cell survival. | | | | | |
| XX | XX | | | | | |
| OS | Homo sapiens. | | | | | |
| XX | XX | | | | | |
| FH | Key | Location/Qualifiers | | | | |
| FT | Peptide | 1..16 | | | | |
| FT | Domain | /label= Signal_peptide | | | | |
| FT | Domain | 17..159 | | | | |
| FT | Domain | /label= C1q_domain_#1 | | | | |
| FT | Domain | 160..328 | | | | |
| FT | Domain | /label= C1q_domain_#2 | | | | |
| XX | XX | | | | | |
| PY | WO200102565-A2. | | | | | |
| XX | XX | | | | | |
| PD | 11-JAN-2001. | | | | | |
| XX | XX | | | | | |
| PF | 28-JUN-2000; 2000WO-US17692. | | | | | |
| XX | XX | | | | | |
| PR | 01-JUL-1999; 99US-0346502. | | | | | |
| XX | XX | | | | | |

PA (ZYMO) ZYMOGENETICS INC.
XX
PI Holloway JL, Lok S;
XX
DR WPI; 2001-138140/14.
DR N-PSDB; AAF28672.
XX
PT Novel secreted protein ZACRP4 polypeptides having tandem C1q globular
PT domains, useful for studying cell-cell communication and regulation of
PT cellular processes -
XX
PS Claim 1; Page 77-78; 82pp; English.
XX
CC The present sequence is human ZACRP4 protein. ZACRP4 protein has two
CC complement factor C1q domains. The ZACRP4 gene is located on human
CC chromosome 11q11. The ZACRP4 coding sequence and protein have a number of
CC uses described in the specification, including, modulation of energy
CC balance and cellular metabolic reactions in mammals. In addition, ZACRP4
CC protein is useful as an autocrine factor, particularly during
CC development, in mediating the processes of an organism, in regulating
CC cellular processes such as cell proliferation and/or differentiation,
CC cell survival and energy balance.
XX
SQ Sequence 329 AA;

Query Match 100.0%; Score 1726; DB 22; Length 329;
Best Local Similarity 100.0%; Pred. No. 1.2e-154;
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLPLLGLLGPACWALGPTPGSGSSELSAFAAARTTPLEGTSEMAVTFDKVYVNIIGD 60
DB 1 MLPLLGLLGPACWALGPTPGSGSSELSAFAAARTTPLEGTSEMAVTFDKVYVNIIGD 60
OY 61 FDVATGQFRCRVPGAYFFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRASQ 120
DB 61 FDVATGQFRCRVPGAYFFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRASQ 120
OY 121 SAMQLDYGDTWVLRHLGAPHYALGAPGATFSGYLVYADADADAPARGPAPPEPRSAFS 180
DB 121 SAMQLDYGDTWVLRHLGAPHYALGAPGATFSGYLVYADADADAPARGPAPPEPRSAFS 180
OY 181 AARTSLVGS DAGPGRHQPLAFDTEFVNIIGD FDDAAGVFRCLPGAYFFSFTLGKLP 240
DB 181 AARTSLVGS DAGPGRHQPLAFDTEFVNIIGD FDDAAGVFRCLPGAYFFSFTLGKLP 240
OY 241 KTLVKLKNRDEVQAMITYDDGASRRREMOSSVMLALRRGDVWLISHDHGCGAYSNH 300
DB 241 KTLVKLKNRDEVQAMITYDDGASRRREMOSSVMLALRRGDVWLISHDHGCGAYSNH 300
OY 301 GKYITFSGFLVYPDLAPAPPGLGASELL 329
DB 301 GKYITFSGFLVYPDLAPAPPGLGASELL 329

RESULT 2
ABG79643
ID ABG79643 standard; Protein; 329 AA.
XX
AC ABG79643;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human novel secreted protein SECP19, Incyte ID No. 931619CD1.
XX
KW Human; SECP; secreted protein; cell proliferative disorder;
KW actinic keratosis; arteriosclerosis; bursitis; hepatitis; cancer;
KW autoimmune disorder; inflammatory disorder; AIDS; asthma; allergy;
KW acquired immunodeficiency syndrome; anaemia; atopic dermatitis;
KW cardiovascular disorder; congestive heart failure; vascular tumour;
KW ischaemic heart disease; myocardial infarction; epilepsy; stroke;
KW hypertensive heart disease; neurological disorder; cerebral neoplasia;
KW Alzheimer's disease; developmental disorder; renal tubular acidosis;
KW Cushing's syndrome; Duchenne muscular dystrophy; hypothyroidism;

KW Becker muscular dystrophy.
XX
OS Homo sapiens.
XX
EN WO200262841-A2.
XX
PD 15-AUG-2002.
XX
PF 28-JAN-2002; 2002WO-US02616.
XX
PR 02-FEB-2001; 2001US-266195P.
PR 08-FEB-2001; 2001US-267924P.
PR 09-FEB-2001; 2001US-267816P.
PR 09-FEB-2001; 2001US-268112P.
PR 26-FEB-2001; 2001US-271639P.
PR 07-SEP-2001; 2001US-317818P.
PR 21-DEC-2001; 2001US-343553P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Tang TY, Yue H, Gandhi AR, Yao MG, Warren BA, Ding L, Dugan BM;
PI Xu Y, Yang J, Thangavelu K, Lal PG, Honcheil CD, Walla NK, Lee S;
PI Lee EA, Richardson TW, Baughn MR, Elliott VS;
XX
XX WPI; 2002-657522/70.
DR N-PSDB; ABS64954.
XX
PT New human secreted proteins and nucleic acids useful in diagnosing,
PT treating and preventing cell proliferative, autoimmune/inflammatory,
PT cardiovascular, neurological, and developmental disorders -
XX
PS Claim 1; Page 140; 158pp; English.

CC The invention relates to twenty four human secreted proteins
CC (SECP1-24), proteins 90% identical to them and active fragments of them.
CC Also included are nucleic acids encoding the SECP proteins, a recombinant
CC polynucleotide comprising a promoter sequence operably linked to the
CC nucleic acid, a cell transformed with the recombinant polynucleotide,
CC a transgenic organism comprising the recombinant polynucleotide, an
CC anti-SECP antibody, and screening for ant/agonists and modulators of
CC SECP function or expression. The SECP proteins and nucleic acids are
CC useful in the diagnosis, treatment and prevention of cell proliferative
CC (e.g. actinic keratosis, arteriosclerosis, bursitis, hepatitis or
CC cancer), autoimmune/inflammatory (e.g. AIDS (acquired immunodeficiency
CC syndrome), asthma, anaemia, allergies or atopic dermatitis),
CC cardiovascular (e.g. congestive heart failure, ischaemic heart disease,
CC myocardial infarction, hypertensive heart disease, or vascular tumours),
CC neurological (e.g. epilepsy, stroke, cerebral neoplasms, or Alzheimer's
CC disease), and developmental (e.g. renal tubular acidosis, Cushing's
CC syndrome, Duchenne and Becker muscular dystrophy, or hypothyroidism)
CC disorders. Many other diseases and disorders are listed in the
CC specification. These may also be used in assessing the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of the secreted proteins. The present sequence represents a
CC SECP protein of the invention.
XX
SQ Sequence 329 AA;

Query Match 98.9%; Score 1707; DB 23; Length 329;
Best Local Similarity 99.4%; Pred. No. 7.6e-153;
Matches 327; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MLPLLGLLGPACWALGPTPGSGSSELSAFAAARTTPLEGTSEMAVTFDKVYVNIIGD 60
DB 1 MLPLLGLLGPACWALGPTPGSGSSELSAFAAARTTPLEGTSEMAVTFDKVYVNIIGD 60
OY 61 FDVATGQFRCRVPGAYFFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRASQ 120
DB 61 FDVATGQFRCRVPGAYFFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRASQ 120
OY 121 SAMQLDYGDTWVLRHLGAPHYALGAPGATFSGYLVYADADADAPARGPAPPEPRSAFS 180
DB 121 SAMQLDYGDTWVLRHLGAPHYALGAPGATFSGYLVYADADADAPARGPAPPEPRSAFS 180

QY 181 AARTSLVGS DAGPGRHQPLAFDTEFVNIGDFDAAAGVFRCLPGAYFFSFTLGKLP R 240
DB 181 AARTSLVGS DAGPGRHQPLAFDTEFVNIGDFDAAAGVFRCLPGAYFFSFTLGKLP R 240
QY 241 KTL SVKLMKNRDEVQAM IYDDGASRRREMOSQSVMLALRRGDVWMLSHDHG YGAYSNH 300
DB 241 KTL SVKLMKNRDEVQAM IYDDGASRRREMOSQSVMLALRRGDVWMLSHDHG YGAYSNH 300
QY 301 GK YITFS GFLVY PDLAPAAPGLGASELL 329
DB 301 GK YITFS GFLVY PDLAPAAPGLGASELL 329

RESULT 3
AAB61424 ID AAB61424 standard; protein; 329 AA.

XX AC AAB61424;
XX DT 04-APR-2001 (first entry)
XX DE Monkey MANGO 245 protein.
XX KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
XX KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
XX KW pancreatic; skeletal; muscle.
XX OS Catarrhini sp..
XX PN WO200100672-A1.
XX PD 04-JAN-2001.
XX PF 29-JUN-2000; 2000WO-US18184.
XX PR 29-JUN-1999; 99US-0342687.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX DR WPI; 2001-050127/06.

PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -

PS Claim 1; Fig 25; 262pp; English.

CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.

XX Sequence 329 AA;

Query Match 96.5%; Score 1666; DB 22; Length 329;
Best Local Similarity 96.7%; Pred. No. 5.7e-149;
Matches 318; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLPLLLGLLPACWALGPTPGSSSELRSAFSAARTTPLEGTSEMAVTFDKVYVNIGD 60
DB 1 MLPLLLGLLPACWALGPAFGSSSELRSAFSAARTTPLEGASEMAVTFDKVYVNIGD 60
QY 61 FDVATGQFRCRVPGAYFFSFTAGKAPHKSLSVMLVNRNDEVQALAFDEQRPGARRASQ 120
DB 61 FDAATGQFRCRVPGAYFFSFTVGKAPHKSLSVMLVNRNDEVQALAFDEQRPSARRASQ 120

QY 121 SAMLQLDYGDVTWLRHLHGAPHYALGAGATFSGYLVVYADADADAPARGPPEPRSAFS 180
DB 121 SAMLQLDYGDVTWLRHLHGAPQYALGAGATFSGYLVVYADADADAPARGPPEPRSAFS 180
QY 181 AARTSLVGS DAGPGRHQPLAFDTEFVNIGDFDAAAGVFRCLPGAYFFSFTLGKLP R 240
DB 181 AARTSLVGS DAGSGPRHRPLAFDTLVNIGDFDAAAGVFRCLPGAYFFSFTLGKLP R 240
QY 241 KTL SVKLMKNRDEVQAM IYDDGASRRREMOSQSVMLALRRGDVWMLSHDHG YGAYSNH 300
DB 241 KTL SVKLMKNRDEVQAM IYDDGASRRREMOSQSVMLALRRGDVWMLSHDHG YGAYSNH 300
QY 301 GK YITFS GFLVY PDLAPAAPGLGASELL 329
DB 301 GK YITFS GFLVY PDLAPAAPGLGAPELL 329

RESULT 4
AAB61423 ID AAB61423 standard; protein; 348 AA.

XX AC AAB61423;
XX DT 04-APR-2001 (first entry)
XX DE Human MANGO 245 protein.
XX KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
XX KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
XX KW pancreatic; skeletal; muscle.
XX OS Homo sapiens.
XX PN WO200100672-A1.
XX PD 04-JAN-2001.
XX PF 29-JUN-2000; 2000WO-US18184.
XX PR 29-JUN-1999; 99US-0342687.
XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX DR WPI; 2001-050127/06.

PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and

PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -

PS Claim 1; Fig 23; 262pp; English.

CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.

XX Sequence 348 AA;

Query Match 90.8%; Score 1567; DB 22; Length 348;
Best Local Similarity 99.7%; Pred. No. 1.4e-139;
Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLPLLLGLLPACWALGPTPGSSSELRSAFSAARTTPLEGTSEMAVTFDKVYVNIGD 60
DB 1 MLPLLLGLLPACWALGPTPGSSSELRSAFSAARTTPLEGTSEMAVTFDKVYVNIGD 60

OY 61 FDVATGQFRCRVPGAYFFSFSTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQ 120
DB 61 FDVATGQFRCRVPGAYFFSFSTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQ 120
OY 121 SAMQLDYGDTVWMLRLHGAPHYALGAPGATFSGYLVVADADADAPARGPPAPPEPRSAFS 180
DB 121 SAMQLDYGDTVWMLRLHGAPHYALGAPGATFSGYLVVADADADAPARGPPAPPEPRSAFS 180
OY 181 AARTSLVGS DAGPGRHQPLAFDTEFVNIGD FDDAAGVRCRLPGAYFFSFFTLGKLP R 240
DB 181 AARTSLVGS DAGPGRHQPLAFDTEFVNIGD FDDAAGVRCRLPGAYFFSFFTLGKLP R 240
OY 241 KTL SVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRRGDVAVMLLSHDHDGYGAYSNH 300
DB 241 KTL SVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRRGDVAVMLLSHDHDGYGAYSNH 300
RESULT 5
ABG70385
ID ABG70385 standard; Protein; 299 AA.
XX
AC ABG70385;
XX
DT 05-NOV-2002 (first entry)
XX
DE Adipocyte complement-related C1q Tumour Necrosis Factor-like protein #3.
XX
KW Human; NOX; NOX-associated disorder; cardiomyopathy; atherosclerosis;
KW cell signal processing; metabolic pathway modulation; metabolic disorder;
KW obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
KW Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
KW haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
KW memory defect; infertility; congenital heart defect; hair growth;
KW pigmentation disorder; endocrine disorder; respiratory disease; health;
KW gastro-intestinal disease; reproductive; neurological disease;
KW bone marrow transplantation; endocrine disease; allergy; inflammation;
KW nephrological disorder; urinary system disorder; age-related disorder;
KW neuropsychiatric disorder; EGF-related protein; SCUBEL; TEN-M4;
KW adipocyte complement-related C1q tumour necrosis factor; out at first;
KW beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;
KW type 1a membrane sushi-containing domain; butyrophillin;
KW type 1a membrane-sushi domain containing.
XX
OS Homo sapiens.
XX
PN WO200257453-A2.
XX
PD 25-JUL-2002.
XX
PF 19-DEC-2001; 2001WO-US50331.
XX
PR 19-DEC-2000; 2000US-265704P.
PR 20-DEC-2000; 2000US-257314P.
PR 02-MAY-2001; 2001US-288153P.
PR 29-MAY-2001; 2001US-294075P.
PR 24-JUL-2001; 2001US-307506P.
PR 10-AUG-2001; 2001US-311590P.
PR 10-AUG-2001; 2001US-311613P.
PR 29-AUG-2001; 2001US-315617P.
PR 14-SEP-2001; 2001US-322358P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Gangoli EA, Patturajan M, Vernet CAM, Maljankar UM, Kekuda R;
PI Stone DJ, Anderson D, Shimkets RA, Burgess CE, Zerhusen BD, Liu X;
PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
DR WPI; 2002-590744/63.
DR N-PSDB; ABBS52097.
XX
PT Novel isolated NOX polypeptide useful for treating cardiomyopathy,
PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious

PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
PT cancer -
XX
PS Claim 1; Page 31; 318pp; English.
XX
CC The present invention relates to new NOX polypeptides. The invention is
CC useful for treating or preventing a NOX-associated disorder such as
CC cardiomyopathy or atherosclerosis, where the disorder is related to cell
CC signal processing and metabolic pathway modulation in a subject,
CC preferably human. The invention is also useful for treating metabolic
CC disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
CC disorders, haematopoietic disorders and various cancers. The molecules of
CC the invention are also useful for treating or preventing cirrhosis,
CC pancreatitis, learning and memory defects, infertility, congenital heart
CC defects, acne, hair growth, pigmentation disorders, endocrine disorders,
CC respiratory disease, gastro-intestinal diseases, reproductive, health,
CC neurological diseases, bone marrow transplantation, endocrine diseases,
CC allergy and inflammation, nephrological disorders, urinary system
CC disorders, neuropsychiatric disorders and age-related disorders.
CC The present amino acid sequence represents a NOX protein of the
CC invention.
XX
SQ Sequence 299. AA;
XX
Query Match 89.2%; Score 1540; DB 23; Length 299;
Best Local Similarity 90.6%; Pred. No. 4e-137;
Matches 298; Conservative 0; Mismatches 1; Indels 30; Gaps 1;
OY 1 MLPLLGLLGPACVNLGPTPGSSSELRSAPSAARTTPLEGTSSEMAVTFDKYVNIIGD 60
DB 1 MLPLLGLLGPACVNLGPTPGSSSELRSAPSAARTTPLEGTSSEMAVTFDKYVNIIGD 60
OY 61 FDVATGQFRCRVPGAYFFSFSTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQ 120
DB 61 FDVATGQFRCRVPGAYFFSFSTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQ 120
OY 121 SAMQLDYGDTVWMLRLHGAPHYALGAPGATFSGYLVVADADADAPARGPPAPPEPRSAFS 180
DB 121 SAMQLDYGDTVWMLRLHGAPHYALGAPGATFSGYLVVADADADAPARGPPAPPEPRSAFS 180
OY 181 AARTSLVGS DAGPGRHQPLAFDTEFVNIGD FDDAAGVRCRLPGAYFFSFFTLGKLP R 240
DB 181 AARTSLVGS DAGPGRHQPLAFDTEFVNIGD FDDAAGVRCRLPGAYFFSFFTLGKLP R 240
OY 241 KTL SVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRRGDVAVMLLSHDHDGYGAYSNH 300
DB 241 KTL SVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRRGDVAVMLLSHDHDGYGAYSNH 300
OY 301 GKYTFSGFLVYPDLAPAAPGLGASELL 329
DB 271 GKYTFSGFLVYPDLAPAAPGLGASELL 299
RESULT 6
AAB61466
ID AAB61466 standard; protein; 334 AA.
XX
AC AAB61466;
XX
DT 04-APR-2001 (first entry)
XX
DE Human MANGO 245 mature protein.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Homo sapiens.
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.

XX 29-JUN-2000; 2000WO-US18184.
PF
XX
XX 29-JUN-1999; 99US-0342687.
PR
XX
XX (MILL-) MILLENNIUM PHARM INC.
PA
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
XX WPI; 2001-050127/06.
DR
XX
XX Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
XX
PS Disclosure; Page 242; 262pp; English.
XX
XX The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
XX Sequence 334 AA;
SQ
Query Match 86.4%; Score 1491; DB 22; Length 334;
Best Local Similarity 99.3%; Pred. No. 2e-132;
Matches 284; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 15 WALGPTPGSGSELRSAPSAARTTPLEGTSEMAVTFDKYVYVNIIGDFDVATGQFRGRVPG 74
Db 1 WALGPTPGSGSELRSAPSAARTTPLEGTSEMAVTFDKYVYVNIIGDFDVATGQFRGRVPG 60
OY 75 AYFFSFTAGKAPHKSLSVMLVRNDEVOALAFDEQRRPGARRAASQSAMLQLDYGDTVWL 134
Db 61 AYFFSFTAGKAPHKSLSVMLVRNDEVOALAFDEQRRPGARRAASQSAMLQLDYGDTVWL 120
OY 135 RLHGAPHYALGAPGATFSGYLVYADADADAPARGPPAPRPSAFSAARTRSLVGSADGP 194
Db 121 RLHGAPQYALGAPGATFSGYLVYADADADAPARGPPAPRPSAFSAARTRSLVGSADGP 180
OY 195 GPRHQPLAFDTEFVNIGDFDAAGVFCRCLPGAYFFSFTLGLPKRTL SVKLMKNRDEV 254
Db 181 GPRHQPLAFDTEFVNIGDFDAAGVFCRCLPGAYFFSFTLGLPKRTL SVKLMKNRDEV 240
OY 255 QAMITYDDGASRRREMOSQSVMLALRRGDAVWL SHDHGAYGAYSNH 300
Db 241 QAMITYDDGASRRREMOSQSVMLALRRGDAVWL SHDHGAYGAYSNH 286
RESULT 7
ABG70384
ID ABG70384 standard; Protein; 284 AA.
XX
AC ABG70384;
XX
DT 05-NOV-2002 (first entry)
XX
DE Adipocyte complement-related C1q Tumour Necrosis Factor-like protein #2.
XX
XX Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis;
KW cell signal processing; metabolic pathway modulation; metabolic disorder;
KW obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
KW Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
KW haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
KW memory defect; infertility; congenital heart defect; hair growth;
KW pigmentation disorder; endocrine disorder; respiratory disease; health;
KW gastro-intestinal disease; reproductive; neurological disease;
KW bone marrow transplantation; endocrine disease; allergy; inflammation;

KW nephrological disorder; urinary system disorder; age-related disorder;
KW neuropsychiatric disorder; EGF-related protein; SCUBE1; TEN-M4;
KW adipocyte complement-related C1q tumour necrosis factor; out at first;
KW beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;
KW type 1a membrane sushi-containing domain; butyrophillin;
XX type 1a membrane-sushi domain containing.
XX
OS Homo sapiens.
XX
XX WO200257453-A2.
PN
XX
XX 25-JUL-2002.
PD
XX
XX 19-DEC-2001; 2001WO-US50331.
PF
XX
XX 19-DEC-2000; 2000US-265704P.
PR 20-DEC-2000; 2000US-257314P.
PR 02-MAY-2001; 2001US-288153P.
PR 29-MAY-2001; 2001US-294075P.
PR 24-JUL-2001; 2001US-307506P.
PR 10-AUG-2001; 2001US-311590P.
PR 10-AUG-2001; 2001US-311613P.
PR 29-AUG-2001; 2001US-315617P.
PR 14-SEP-2001; 2001US-322358P.
XX
XX (CURA-) CURAGEN CORP.
PA
XX
XX Gangolli EA, Paturajan M, Vernet CAM, Malyankar UM, Kekuda R;
PI Stone DJ, Anderson D, Shimkets RA, Burgess CE, Zerhusen BD, Liu X;
PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
XX WPI; 2002-590744/63.
DR N-PSDB; ABS52096.
DR
XX
XX Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious
PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
PT cancer -
XX
XX Claim 1; Page 30; 318pp; English.
PS
XX
XX The present invention relates to new NOVX polypeptides. The invention is
CC useful for treating or preventing a NOVX-associated disorder such as
CC cardiomyopathy or atherosclerosis, where the disorder is related to cell
CC signal processing and metabolic pathway modulation in a subject,
CC preferably human. The invention is also useful for treating metabolic
CC disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
CC disorders, haematopoietic disorders and various cancers. The molecules of
CC the invention are also useful for treating or preventing cirrhosis,
CC pancreatitis, learning and memory defects, infertility, congenital heart
CC defects, acne, hair growth, pigmentation disorders, endocrine disorders,
CC respiratory disease, gastro-intestinal diseases, reproductive health,
CC neurological diseases, bone marrow transplantation, endocrine diseases,
CC allergy and inflammation, nephrological disorders, urinary system
CC disorders, neuropsychiatric disorders and age-related disorders.
CC The present amino acid sequence represents a NOVX protein of the
CC invention.
XX
XX Sequence 284 AA;
SQ
Query Match 83.7%; Score 1445.5; DB 23; Length 284;
Best Local Similarity 86.0%; Pred. No. 3.2e-128;
Matches 283; Conservative 0; Mismatches 1; Indels 45; Gaps 1;
OY 1 MLPLLLGLLGPACWALGPTPGSGSELRSAPSAARTTPLEGTSEMAVTFDKYVYVNIIGD 60
Db 1 MLPLLLGLLGPACWALGPTPGSGSELRSAPSAARTTPLEGTSEMAVTFDKYVYVNIIGD 60
OY 61 FDVATGQFRGRVPGAYFFSFTAGKAPHKSLSVMLVRNDEVOALAFDEQRRPGARRAASQ 120
Db 61 FDVATGQFRGRVPGAYFFSFTAGKAPHKSLSVMLVRNDEVOALAFDEQRRPGARRAASQ 120

QY 121 SAMQLDYGDTWMLRLHGAPHYALGAPGATFSGYLVYADADADAPARGPPAPPEPRSAFS 180
Db 121 SAMQLDYGDTWMLRLHGAPHYALGAPGATFSGYLVYADA-----160
QY 181 AARTRSLVGS DAGPGPRHQPLAFDTEFVNIGGDFDAAGVFRCRLLPGAYFFSFTLGKLP 240
Db 161 -----EFVNIGGDFDAAGVFRCRLLPGAYFFSFTLGKLP 195
QY 241 KTL SVKLMKNRDEVQAM IYDDGASRRREMOSQSVMLALRRGDVWLLSHDHGYGAYSNH 300
Db 196 KTL SVKLMKNRDEVQAM IYDDGASRRREMOSQSVMLALRRGDVWLLSHDHGYGAYSNH 255
QY 301 GK YITFSGFLVYPDLAPAAPPGLGASELL 329
Db 256 GK YITFSGFLVYPDLAPAAPPGLGASELL 284

RESULT 8
AAB61488
ID AAB61488 standard; protein; 199 AA.
XX AAB61488;
AC
XX
DT 04-APR-2001 (first entry)
XX
DE Murine MANGO 245 protein.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Mus musculus.
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Claim 1; Fig 29; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 199 AA;

Query Match 51.1%; Score 882; DB 22; Length 199;
Best Local Similarity 87.2%; Pred. No. 3.5e-75;
Matches 177; Conservative 4; Mismatches 18; Indels 4; Gaps 3;

QY 1 MLPLLLGLGPAA CWALGPTPGSGSSELRSAFSAARTTPLEGTSEMAVTFDKVYVNIGD 60
Db 1 MLTLLGLGPAA CWALGPA-GPGSSELRSAFSAARTTPLEGTSEMAVTFDKVYVNIGD 59

QY 61 FDVATGQFRCRVPGAVFESFTAGAKPHKSLSVMLVRNDEVQALAFDEQRRPGARRASQ 120
Db 60 FDAATGRFRCRVPGAVFESFTAGAKPHKNSVMLVRNDEVQALAFDKORRPGARRASQ 119
QY 121 SAMQLDYGDTWMLRLHGAPHYALGAPGATFSGYLVYADADADAPARGPPAPPEPRSAFS 180
Db 120 SAMQLDYGDTWMLRLHGAPHYALGAPGATFSGYLVYADADADAPARG-PAAPPEPRSAFS 178
QY 181 AARTRSLVGS DAGPGPRHQPLAF 203
Db 179 ARHATWAPNPPRPAPRR--LAF 199

RESULT 9
AAB61479
ID AAB61479 standard; protein; 192 AA.
XX AAB61479;
AC
XX
DT 04-APR-2001 (first entry)
XX
DE Human MANGO 245 cytoplasmic domain.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Homo sapiens.
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Disclosure; Page 247; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 192 AA;

Query Match 44.0%; Score 760; DB 22; Length 192;
Best Local Similarity 100.0%; Pred. No. 1.1e-63;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 YADADADAPARGPPAPPEPRSAFSAARTRSLVGS DAGPGPRHQPLAFDTEFVNIGGDFDA 216
Db 1 YADADADAPARGPPAPPEPRSAFSAARTRSLVGS DAGPGPRHQPLAFDTEFVNIGGDFDA 60
QY 217 AAGVFRCRLLPGAYFFSFTLGKLPKRLTSVKLMKNRDEVQAM IYDDGASRRREMOSQSVML 276
Db 61 AAGVFRCRLLPGAYFFSFTLGKLPKRLTSVKLMKNRDEVQAM IYDDGASRRREMOSQSVML 120

QY 277 ALRRGDVWLLSHDHGCGAYSNNH 300
Db 121 ALRRGDVWLLSHDHGCGAYSNNH 144
RESULT 10
ID ABG70383 standard; Protein; 221 AA.
XX
AC ABG70383;
DT 05-NOV-2002 (first entry)
XX
DE Adipocyte complement-related C1q Tumour Necrosis Factor-like protein #1.
XX
KM Human; NOX; NOX-associated disorder; cardiomyopathy; atherosclerosis;
KM cell signal processing; metabolic pathway modulation; metabolic disorder;
KM obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
KM Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
KM haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
KM memory defect; infertility; congenital heart defect; hair growth;
KM pigmentation disorder; endocrine disorder; respiratory disease; health;
KM gastro-intestinal disease; reproductive; neurological disease;
KM bone marrow transplantation; endocrine disease; allergy; inflammation;
KM nephrological disorder; urinary system disorder; age-related disorder;
KM neuropsychiatric disorder; EGF-related protein; SCUB1; TEN-M4;
KM adipocyte complement-related C1q tumour necrosis factor; out at first;
KM beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;
KM type 1a membrane sushi-containing domain; butyrophillin;
KM type 1a membrane-sushi domain containing.
XX
OS Homo sapiens.
XX
PN WO200257453-A2.
XX
PD 25-JUL-2002.
XX
PF 19-DEC-2001; 2001WO-US50331.
XX
PR 19-DEC-2000; 2000US-265704P.
PR 20-DEC-2000; 2000US-257314P.
PR 02-MAY-2001; 2001US-288153P.
PR 29-MAY-2001; 2001US-294075P.
PR 24-JUL-2001; 2001US-307506P.
PR 10-AUG-2001; 2001US-311590P.
PR 10-AUG-2001; 2001US-311613P.
PR 29-AUG-2001; 2001US-315617P.
PR 14-SEP-2001; 2001US-322358P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Gangoli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
PI Stone DJ, Anderson D, Shinkets RA, Burgess CE, Zerhusen BD, Liu X;
PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
XX WPI; 2002-590744/63.
DR N-PSDB; ABS52095.
XX
PT Novel isolated NOX polypeptide useful for treating cardiomyopathy,
PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious
PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
PT cancer.
XX
XX
PS Claim 1; Page 28; 318pp; English.
XX
CC The present invention relates to new NOX polypeptides. The invention is
CC useful for treating or preventing a NOX-associated disorder such as
CC cardiomyopathy or atherosclerosis, where the disorder is related to cell
CC signal processing and metabolic pathway modulation in a subject,
CC preferably human. The invention is also useful for treating metabolic
CC disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
CC disorders, haematopoietic disorders and various cancers. The molecules of

CC the invention are also useful for treating or preventing cirrhosis,
CC pancreatitis, learning and memory defects, infertility, congenital heart
CC defects, acne, hair growth, pigmentation disorders, endocrine disorders,
CC respiratory disease, gastro-intestinal diseases, reproductive, health,
CC neurological diseases, bone marrow transplantation, endocrine diseases,
CC allergy and inflammation, nephrological disorders, urinary system
CC disorders, neuropsychiatric disorders and age-related disorders.
CC The present amino acid sequence represents a NOX protein of the
CC invention.
XX
SQ Sequence 221 AA;
XX
Query Match 41.0%; Score 707.5; DB 23; Length 221;
Best Local Similarity 49.5%; Pred. No. 1.3e-58;
Matches 153; Conservative 1; Mismatches 2; Indels 153; Gaps 1;
QY 21 PGPSSSELRSAFSAARTTPLEGTSEMAVTFDKYVNIIGDFDVATGQFRCPGAYFFSF 80
Db 66 PGDSSSAMRSAFSAARTTPLEGTSEMAVTFDKY----- 99
QY 81 TAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQAMQLDYGDVWLRLLHGAP 140
Db 100 ----- 99
QY 141 HYALGAPGATPGSYLVYADADADAPARCPAPPEPRSAFSAARTSLVGS DAGPGRHQP 200
Db 100 ----- 99
QY 201 LAFDTEFVNIIGDFDAAGVFCRLPGAYFPSTLGLPRKTL SVKLMKNRDEVQAMITYD 260
Db 100 -----VNIGDFDAAGVFCRLPGAYFPSTLGLPRKTL SVKLMKNRDEVQAMITYD 152
QY 261 DGASRRREMOSQSVMLALRRGDVWLLSHDHGCGAYSNNHGKXITFSGFLVYPDLAPAP 320
Db 153 DGASRRREMOSQSVMLALRRGDVWLLSHDHGCGAYSNNHGKXITFSGFLVYPDLAPAP 212
QY 321 PGLGASELL 329
Db 213 PGLGASELL 221
RESULT 11
AAB61473
ID AAB61473 standard; protein; 134 AA.
XX
AC AAB61473;
XX
DT 04-APR-2001 (first entry)
XX
DE Monkey MANGO 245 C1q domain #2.
XX
KM TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KM autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KM pancreatic; skeletal; muscle.
XX
OS Catarrhini sp..
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases

| | |
|-----------|--|
| PT | (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver disorders (e.g. jaundice) - |
| PS | Disclosure; Fig 28; 262pp; English. |
| XX | |
| CC | The present invention relates to cDNAs encoding TANGO 244. |
| CC | TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins. |
| CC | The nucleic acids, proteins and protein modulators are useful for treating colonic disorders, inflammatory diseases, tumors, |
| CC | renal disorders, liver disorders, lung disorders, autoimmune diseases, allergic diseases, cardiovascular diseases, brain disorders, |
| CC | degenerative diseases placental, pancreatic, skeletal and muscle disorders. |
| CC | |
| XX | |
| SQ | Sequence 134 AA; |
| | |
| | Query Match 39.6%; Score 683; DB 22; Length 134; |
| | Best Local Similarity 97.8%; Pred. No. 1.3e-56; |
| | Matches 131; Conservative 1; Mismatches 2; Indels 0; Gaps 0 |
| Oy | 178 AFSARTSLVSGSDAGPGPRHQPLAFDTEFVNIGDFDAAGVFRCLPGAYFFSFTLGK 237 1 AFSAARTSLVSGSDAGSGPRHRPLAFDTLVNIGDFDAAGVFRCLPGAYFFSFTLGK 60 |
| Oy | 238 LPRKTLVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRRGDVWLLSHDHGYGAY 297 61 LPKRITLTKMKNRDEVQAMITYDDGASRRREMOSQSVMLALRRGDVWLLSHDHGYGAY 120 |
| Oy | 298 SNHGKYITFSGFLV 311 121 SNHGKYITFSGFLV 134 |
| Dd | |
| | |
| RESULT 12 | |
| ABB11578 | |
| ID | ABB11578 standard; peptide; 127 AA. |
| XX | |
| AC | ABB11578; |
| XX | |
| DT | 11-JAN-2002 (first entry) |
| DE | |
| XX | Human secreted protein homologue, SEQ ID NO:1948. |
| XX | |
| KW | Human; cytokine; cell proliferation; cell differentiation; growth factor; haematopoiesis regulation; tissue growth; immunomodulator; activin; inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis; proliferation; metastasis; cancer; tumour; haematopoietic disorder; myeloid cell disorder; lymphoid cell disorder; asthma; arthritis; chronic inflammatory condition; proliferative retinopathy; atherosclerosis; coronary heart disease; arterial ischaemia; bone disorder; osteoporosis; vascular growth disorder; tissue regeneration; wound healing; infection; immune disorder; cell culture; drug screening; gene therapy; anti-inflammatory; antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic; cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial; antifungal; vulnerary; antitumor. |
| KW | |
| XX | |
| OS | Homo sapiens. |
| XX | |
| PN | WO200157188-A2. |
| XX | |
| PD | 09-AUG-2001. |
| XX | |
| PF | 05-FEB-2001; 2001WO-US03800. |
| XX | |
| PR | 03-FEB-2000; 2000US-0496914. |
| XX | |
| PA | 27-APR-2000; 2000US-0560875. |
| XX | |
| PI | (HYSE-) HYSEQ INC. |
| XX | |
| PI | Tang YT, Liu C, Drmanac RT, |
| XX | |
| DR | WPI; 2001-457740/49. |

| | |
|-----------|--|
| DR | N-PSDB; ABA08822. |
| XX | |
| PT | Human proteins and DNA encoding sequences useful for preventing, |
| PT | treating or ameliorating a medical condition in a mammalian subject |
| PT | e.g. arthritis and cancer - |
| XX | |
| PS | Claim 20; Page 214-215; 1963pp; English. |
| XX | |
| CC | Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and |
| CC | sequences ABA08225-ABA09574 represent nucleic acids encoding them. The |
| CC | invention also relates to vectors and recombinant host cells comprising a |
| CC | nucleotide of the invention, methods of producing the novel polypeptides, |
| CC | antibodies against the polypeptides, methods of detecting the nucleotides |
| CC | or polypeptides in a sample, and methods of identifying compounds which |
| CC | bind to polypeptides of the invention. Although novel, many of the |
| CC | polypeptides of the invention have homology to known proteins, thereby |
| CC | providing an insight into their probable biological activities, and hence |
| CC | potential therapeutic applications. The polypeptides of the invention may |
| CC | have various activities, including cytokine, cell proliferation or cell |
| CC | differentiation activities; stem cell growth factor activity; |
| CC | haematopoiesis regulatory activity; tissue growth activity; |
| CC | immunomodulatory activity; activin- or inhibin-related activities; |
| CC | chemotactic or chemokinetic activities; haemostatic, thrombotic or |
| CC | thrombolytic activities; receptor or ligand activities; or may be |
| CC | involved in oncogenesis, cancer cell proliferation or metastasis. |
| CC | Depending on their biological activities, polypeptides and nucleotides of |
| CC | the invention are useful for preventing, treating or ameliorating medical |
| CC | conditions, e.g., by protein or gene therapy. Such conditions include |
| CC | cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell |
| CC | dysorders), chronic inflammatory conditions (e.g., asthma or arthritis), |
| CC | proliferative retinopathy, atherosclerosis, coronary heart disease, |
| CC | arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal |
| CC | vascular growth. Polypeptides involved with tissue regeneration and |
| CC | repair (or nucleic acids encoding them) may be used to promote wound |
| CC | healing (e.g., of burns, incisions and ulcers), while those with |
| CC | immunomodulatory activities may be used in the treatment of viral, |
| CC | bacterial and fungal infections in addition to immune disorders. |
| CC | Polypeptides with growth factor activity may be used in cell cultures to |
| CC | promote cell growth. For example, such polypeptides may be used to |
| CC | manipulate stem cells in culture to give rise to neuroepithelial cells, |
| CC | that can be used to augment or replace cells damaged by illness, |
| CC | autoimmune disease or accidental damage. The polypeptides and nucleotides |
| CC | may also be used in the diagnosis of the above conditions, and in drug |
| CC | screening techniques. The present sequence represents a novel human |
| CC | polypeptide of the invention. |
| CC | |
| XX | |
| SQ | Sequence 127 AA; |
| | |
| | Query Match 38.5%; Score 664; DB 22; Length 127; |
| | Best Local Similarity 99.2%; Pred. No. 7.8e-55; |
| | Matches 126; Conservative 1; Mismatches 0; Indels 0; Gaps 0; |
| OY | 203 FDTFVNIGGDFDAAAGVFRCLPGAVFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDG 262 1 FDTFVNIGGDFDAAAGVFRCLPGAVFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDG 60 |
| OY | 263 ASRRREMOSQSVMLALRRGDVAWVLISHDHGYGAYSNHGXITTFSGFLVYPDLAPAAPG 322 : 61 SSRREMOSQSVMLALRRGDVAWVLISHDHGYGAYSNHGXITTFSGFLVYPDLAPAAPG 120 |
| Db | 323 LGASELL 329 121 LGASELL 127 |
| Db | |
| RESULT 13 | |
| AAB61477 | |
| ID | AAB61477 standard; protein; 125 AA. |
| AC | AAB61477; |
| DT | 04-APR-2001 (first entry) |
| XX | |

DE Human MANGO 245 extracellular domain.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Homo sapiens.
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Disclosure; Page 246; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 125 AA;

Query Match 37.7%; Score 650; DB 22; Length 125;
Best Local Similarity 100.0%; Pred. No. 1.6e-53;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 WALGPTPGSGSELRSAFSAARTTPLEGTESEMAVTFDKYVYVNIIGDFDVATGQFRCPVG 74
Db 1 WALGPTPGSGSELRSAFSAARTTPLEGTESEMAVTFDKYVYVNIIGDFDVATGQFRCPVG 60

QY 75 AYFESFTAGKAPHKSLSVMLVNRNDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTVWL 134
Db 61 AYFESFTAGKAPHKSLSVMLVNRNDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTVWL 120

QY 135 RLHGA 139
Db 121 RLHGA 125

RESULT 14
AAB61469
ID AAB61469 standard; protein; 126 AA.
XX
AC AAB61469;
XX
DT 04-APR-2001 (first entry)
XX
DE Human MANGO 245 Clq domain #1.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Homo sapiens.
XX

PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Disclosure; Fig 26; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 126 AA;

Query Match 37.4%; Score 646; DB 22; Length 126;
Best Local Similarity 99.2%; Pred. No. 3.9e-53;
Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 31 AFSARTTPLEGTESEMAVTFDKYVYVNIIGDFDVATGQFRCPVGAYFFSFTAGKAPHKSL 90
Db 1 AFSARTTPLEGTESEMAVTFDKYVYVNIIGDFDVATGQFRCPVGAYFFSFTAGKAPHKSL 60

QY 91 SVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTVWLRLHGAPHYALGAPGAT 150
Db 61 SVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTVWLRLHGAPHYALGAPGAT 120

QY 151 FSGYLV 156
Db 121 FSGYLV 126

RESULT 15
AAB61489
ID AAB61489 standard; protein; 126 AA.
XX
AC AAB61489;
XX
DT 04-APR-2001 (first entry)
XX
DE Murine MANGO 245 Clq domain.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Mus musculus.
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX

PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Disclosure; Page 260-261; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 126 AA;

Query Match 37.0%; Score 639; DB 22; Length 126;
Best Local Similarity 96.8%; Pred. No. 1.8e-52;
Matches 122; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 31 AFSARTTPECTSEMAVTFDKVYVNIIGDFDVATGQFRRCRVPGAYFFSFTAGKAPHKSL 90
|||
Db 1 AFSARTTPECTSEMAVTFDKVYVNIIGDFDAGTGRFRCRVPGAYFFSFTAGKAPHKSL 60

QY 91 SVMLVNRNDEVQALAFDEQRRPGARRAASQSAAMLQLDYGDTVWLRHGAAPHYALGAPGAT 150
|||
Db 61 SVMLVNRNDEVQALAFDQRRPGARRAASQSAAMLQLDYGDTVWLRHGAAPHYALGAPGAT 120

QY 151 FSGYLV 156
|||
Db 121 FSGYLV 126

Search completed: January 12, 2004, 08:15:34
Job time : 51.5798 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:15:44 ; Search time 40.0579 Seconds
(without alignments)
1655.023 Million cell updates/sec

Title: US-10-085-167-2
Perfect score: 1726
Sequence: 1 MLPLLLGLLGPACWALGPT.....LVYPDLAPAPPGLGASELL 329

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747907 seqs, 201509753 residues

Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------------|-------------------|
| 1 | 1726 | 100.0 | 329 | US-10-085-167-2 | Sequence 2, Appli |
| 2 | 1720 | 99.7 | 329 | US-10-236-055A-14 | Sequence 14, Appl |
| 3 | 1587 | 91.9 | 326 | US-10-236-055A-16 | Sequence 16, Appl |
| 4 | 237.5 | 13.8 | 252 | US-09-893-737-52 | Sequence 52, Appl |
| 5 | 230 | 13.3 | 202 | US-10-091-458-43 | Sequence 43, Appl |
| 6 | 230 | 13.3 | 420 | US-10-236-055A-2 | Sequence 2, Appli |
| 7 | 230 | 13.3 | 746 | US-09-738-973-185 | Sequence 185, App |
| 8 | 230 | 13.3 | 746 | US-09-854-133-185 | Sequence 185, App |
| 9 | 230 | 13.3 | 746 | US-10-144-649A-185 | Sequence 185, App |
| 10 | 230 | 13.3 | 908 | US-10-309-422-10 | Sequence 10, Appl |
| 11 | 230 | 13.3 | 909 | US-10-309-422-22 | Sequence 22, Appl |
| 12 | 230 | 13.3 | 957 | US-10-309-422-14 | Sequence 14, Appl |
| 13 | 230 | 13.3 | 958 | US-10-309-422-26 | Sequence 26, Appl |
| 14 | 230 | 13.3 | 992 | US-10-309-422-18 | Sequence 18, Appl |
| 15 | 230 | 13.3 | 993 | US-10-309-422-30 | Sequence 30, Appl |

| | | | | | | |
|----|-------|------|------|----|-------------------|-------------------|
| 16 | 230 | 13.3 | 1042 | 12 | US-10-309-422-8 | Sequence 8, Appli |
| 17 | 230 | 13.3 | 1043 | 12 | US-10-309-422-20 | Sequence 20, Appl |
| 18 | 230 | 13.3 | 1091 | 12 | US-10-309-422-12 | Sequence 12, Appl |
| 19 | 230 | 13.3 | 1092 | 12 | US-10-309-422-24 | Sequence 24, Appl |
| 20 | 230 | 13.3 | 1126 | 12 | US-10-309-422-16 | Sequence 16, Appl |
| 21 | 230 | 13.3 | 1127 | 12 | US-10-309-422-28 | Sequence 28, Appl |
| 22 | 226 | 13.1 | 409 | 12 | US-10-236-055A-4 | Sequence 4, Appli |
| 23 | 220.5 | 12.8 | 281 | 10 | US-09-911-176B-44 | Sequence 44, Appl |
| 24 | 220.5 | 12.8 | 281 | 12 | US-10-360-186-44 | Sequence 44, Appl |
| 25 | 220.5 | 12.8 | 281 | 12 | US-10-392-531-44 | Sequence 44, Appl |
| 26 | 220.5 | 12.8 | 281 | 12 | US-10-392-706-44 | Sequence 44, Appl |
| 27 | 220.5 | 12.8 | 281 | 12 | US-10-385-015-5 | Sequence 5, Appli |
| 28 | 220.5 | 12.8 | 281 | 15 | US-10-180-762-44 | Sequence 44, Appl |
| 29 | 220.5 | 12.8 | 281 | 15 | US-10-241-258-44 | Sequence 44, Appl |
| 30 | 216 | 12.5 | 281 | 9 | US-09-768-826-46 | Sequence 46, Appl |
| 31 | 215 | 12.5 | 281 | 9 | US-09-765-205-24 | Sequence 24, Appl |
| 32 | 215 | 12.5 | 281 | 9 | US-09-866-028-78 | Sequence 78, Appl |
| 33 | 215 | 12.5 | 281 | 10 | US-09-944-449-78 | Sequence 78, Appl |
| 34 | 215 | 12.5 | 281 | 10 | US-09-944-457-78 | Sequence 78, Appl |
| 35 | 215 | 12.5 | 281 | 10 | US-09-944-862-78 | Sequence 78, Appl |
| 36 | 215 | 12.5 | 281 | 10 | US-09-945-587-78 | Sequence 78, Appl |
| 37 | 215 | 12.5 | 281 | 10 | US-09-945-015-78 | Sequence 78, Appl |
| 38 | 215 | 12.5 | 281 | 10 | US-09-944-396-78 | Sequence 78, Appl |
| 39 | 215 | 12.5 | 281 | 10 | US-09-944-097-78 | Sequence 78, Appl |
| 40 | 215 | 12.5 | 281 | 10 | US-09-944-432-78 | Sequence 78, Appl |
| 41 | 215 | 12.5 | 281 | 10 | US-09-943-762-78 | Sequence 78, Appl |
| 42 | 215 | 12.5 | 281 | 10 | US-09-944-654-78 | Sequence 78, Appl |
| 43 | 215 | 12.5 | 281 | 10 | US-09-943-851A-78 | Sequence 78, Appl |
| 44 | 215 | 12.5 | 281 | 10 | US-09-944-413-78 | Sequence 78, Appl |
| 45 | 215 | 12.5 | 281 | 10 | US-09-911-176B-2 | Sequence 2, Appli |

ALIGNMENTS

RESULT 1
US-10-085-167-2
; Sequence 2, Application US/10085167
; Publication No. US20030170781A1
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Lok, Si
; TITLE OF INVENTION: SECRETED PROTEIN ZACRP4
; FILE REFERENCE: 99-29
; CURRENT APPLICATION NUMBER: US/10/085,167
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: 60/141,928
; PRIOR FILING DATE: 1999-07-01
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 2
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-085-167-2

| | | | | |
|-----------------------|-----------------|----------------------------------|---------------------------------|-------------|
| Query Match | 100.0%; | Score 1726; | DB 12; | Length 329; |
| Best Local Similarity | 100.0%; | Pred. No. 3.4e-154; | | |
| Matches 329; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| QY | 1 | MLPLLLGLLGPACWALGPTPGGSSELRSASAA | RTTPLEGTSEMAVTFDKYVNI | IGD 60 |
| Db | 1 | MLPLLLGLLGPACWALGPTPGGSSELRSASAA | RTTPLEGTSEMAVTFDKYVNI | IGD 60 |
| QY | 61 | FDVATGQFRCRYGAYFFSFTAGKAPHKSL | SVMLVRNRDEVQALAFDEQRRPGARRASQ | 120 |
| Db | 61 | FDVATGQFRCRYGAYFFSFTAGKAPHKSL | SVMLVRNRDEVQALAFDEQRRPGARRASQ | 120 |
| QY | 121 | SAMLQLDYGDYTWLRLHGAPHYALGAPGAT | FTSGYLVYADADADAPARGPPAPPEPRSAFS | 180 |
| Db | 121 | SAMLQLDYGDYTWLRLHGAPHYALGAPGAT | FTSGYLVYADADADAPARGPPAPPEPRSAFS | 180 |
| QY | 181 | AARTRSLVGSADGPGPRHQPLAFDTFVN | IGGDFDAAGVFRRCRLPGAYFFSFTLGKLP | 240 |

Db 181 AARTSLVGS DAGPGRHQPLAFDTEFVNIGGDFDAAGVFRCLPGAYFFSFTLGKLP 240
QY 241 KTL SVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRRGDAVWLLSHDHGYGAYSNNH 300
Db 241 KTL SVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRRGDAVWLLSHDHGYGAYSNNH 300
QY 301 GK YITFSGFLVYPDLAPAAPPGLGASELL 329
Db 301 GK YITFSGFLVYPDLAPAAPPGLGASELL 329

RESULT 2

US-10-236-055A-14
; Sequence 14, Application US/10236055A
; Publication No. US20030134328A1
; GENERAL INFORMATION:
; APPLICANT: Basham, Beth E.
; APPLICANT: Forsythe, Ian
; APPLICANT: Gorman, Daniel M.
; APPLICANT: Mattson, Jeanine
; APPLICANT: Moshrefi, Mehrdad
; APPLICANT: Parham, Christi
; TITLE OF INVENTION: MAMMALIAN GENES; RELATED REAGENTS
; FILE REFERENCE: DX01343K
; CURRENT APPLICATION NUMBER: US/10/236, 055A
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/317, 988
; PRIOR FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-236-055A-14

Query Match 99.7%; Score 1720; DB 12; Length 329;
Best Local Similarity 99.7%; Pred. No. 1.3e-153;
Matches 328; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MLPLLLGLGPACWALGPTPGSSSELRSASFSAARTTPLEGTSEMAVTFDKVYVNIIGD 60
Db 1 MLPLLLGLGPACWALGPTPGSSSELRSASFSAARTTPLEGTSEMAVTFDKVYVNIIGD 60
QY 61 FDVATGQFRRCRVPGAYFFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRASQ 120
Db 61 FDVATGQFRRCRVPGAYFFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRASQ 120
QY 121 SAMLQLDYGDVTWMLRLHGAPHYALGAPGATFSGYLVYADADADAPARGPPAPPEPRSAFS 180
Db 121 SAMLQLDYGDVTWMLRLHGAPHYALGAPGATFSGYLVYADADADAPARGPPAPPEPRSAFS 180
QY 181 AARTSLVGS DAGPGRHQPLAFDTEFVNIGGDFDAAGVFRCLPGAYFFSFTLGKLP 240
Db 181 AARTSLVGS DAGPGRHQPLAFDTEFVNIGGDFDAAGVFRCLPGAYFFSFTLGKLP 240
QY 241 KTL SVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRRGDAVWLLSHDHGYGAYSNNH 300
Db 241 KTL SVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRRGDAVWLLSHDHGYGAYSNNH 300
QY 301 GK YITFSGFLVYPDLAPAAPPGLGASELL 329
Db 301 GK YITFSGFLVYPDLAPAAPPGLGASELL 329

RESULT 3

US-10-236-055A-16
; Sequence 16, Application US/10236055A
; Publication No. US20030134328A1
; GENERAL INFORMATION:
; APPLICANT: Basham, Beth E.
; APPLICANT: Forsythe, Ian

; APPLICANT: Gorman, Daniel M.
; APPLICANT: Mattson, Jeanine
; APPLICANT: Moshrefi, Mehrdad
; APPLICANT: Parham, Christi
; TITLE OF INVENTION: MAMMALIAN GENES; RELATED REAGENTS
; FILE REFERENCE: DX01343K
; CURRENT APPLICATION NUMBER: US/10/236, 055A
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/317, 988
; PRIOR FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-236-055A-16
Query Match 91.9%; Score 1587; DB 12; Length 326;
Best Local Similarity 94.2%; Pred. No. 4.2e-141;
Matches 309; Conservative 2; Mismatches 15; Indels 2; Gaps 2;

QY 1 MLPLLLGLGPACWALGPTPGSSSELRSASFSAARTTPLEGTSEMAVTFDKVYVNIIGD 60
Db 1 MLPLLLGLGPACWALGPA-GPGSSSELRSASFSAARTTPLEGTSEMAVTFDKVYVNIIGD 59
QY 61 FDVATGQFRRCRVPGAYFFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRASQ 120
Db 60 FDAATGRFRCRVPGAYFFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRASQ 119
QY 121 SAMLQLDYGDVTWMLRLHGAPHYALGAPGATFSGYLVYADADADAPARGPPAPPEPRSAFS 180
Db 120 SAMLQLDYGDVTWMLRLHGAPHYALGAPGATFSGYLVYADADADAPARG-PAAPPEPRSAFS 178
QY 181 AARTSLVGS DAGPGRHQPLAFDTEFVNIGGDFDAAGVFRCLPGAYFFSFTLGKLP 240
Db 179 AARTSLVGS DAGPGRHRPLAFDTELVNIGGDFDAAGVFRCLPGAYFFSFTLGKLP 238
QY 241 KTL SVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRRGDAVWLLSHDHGYGAYSNNH 300
Db 239 KTL SVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLPLRRGDAVWLLSHDHGYGAYSNNH 298
QY 301 GK YITFSGFLVYPDLAPAAPPGLGASEL 328
Db 299 GK YITFSGFLVYPDLAAGPPALKPPEL 326

RESULT 4

US-09-893-737-52
; Sequence 52, Application US/09893737
; Patent No. US20020110855A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell, Scott R.
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
; FILE REFERENCE: 00-41
; CURRENT APPLICATION NUMBER: US/09/893, 737
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215, 446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-893-737-52

Query Match 13.8%; Score 237.5; DB 10; Length 252;
Best Local Similarity 43.0%; Pred. No. 3.7e-14;
Matches 65; Conservative 12; Mismatches 61; Indels 13; Gaps 7;
QY 168 GPPAPPEPR--SAFSAARTSLVGS DAGPGRHQPLAFDTEFVNIGGDFDAAGVFRCL 225

```
Db      106 GPGGACRYYAFAFVGREGHSSD-----HFQAVPFDTLVLNDGAFDLAAGRELCIV 160
QY      226 PGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRGDAVW 285
Db      161 PGVYFLSLNVHTWYKETYLHMLNR-RPAVLYAQ-PSERSVMQAQSLMLLLAAGDAVW 218
QY      286 --LISHDHGCGAYSNGK-YITFSGFLVYP 313
Db      219 VRMFQDRDN-AIYGEHGDLYITFSGHLVKP 248

RESULT 5
US-10-091-458-43
; Sequence 43, Application US/10091458
; Publication No. US20030068627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ06C1
; CURRENT APPLICATION NUMBER: US/10/091,458
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 09/764,900
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
```

PRIOR APPLICATION NUMBER: 60/235,836
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/230,438
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215,135
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/225,266
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/249,218
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,208
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,213
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,212
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,207
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,245
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,244
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,217
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,211
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,215
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,264
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,214
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,297
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/232,400
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/231,242
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,081
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,080
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,414
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,244
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,064
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/233,063
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,397
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,399
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,401
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/241,808
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,826
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,786
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,221
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,475
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/231,243
PRIOR FILING DATE: 2000-09-08

Query Match 13.3%; Score 230; DB 15; Length 202;
Best Local Similarity 38.4%; Pred. No. 1.4e-13;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

QY 163 DAPARGP-----PAPPEPSAFASTRSLVSGDAGPGRHOPLAFDTEFVNIGG 212
DB 49 DVPVTNPAATILPVHYVPLPQOMRVAFAAART-----SNLAPGTLDPVFDLLNNIGE 103
QY 213 DFDAAAGVFCRLPGAYFFSFTLGKLPR-KRTLSVKLMKNRDEVOAMITYDDGASRRREMOS 271
DB 104 TFDLQGRFNCVPVNGTYVFIHMLKLA VNPPLYVNLKNEEVLVSAYANDGAP-DHETAS 162
QY 272 QSVMLALRGGDAVWLISHDHDGYGA-YSNHGKYITFSGFLVYPD 314
DB 163 NHAITQLFGQDQIWLRLH---RGAIYSSWKXSTFSGYLLYQD 202

RESULT 6
US-10-236-055A-2
Sequence 2, Application US/102336055A
Publication No. US20030134328A1
GENERAL INFORMATION:
APPLICANT: Basham, Beth E.
APPLICANT: Forsythe, Ian
APPLICANT: Gorman, Daniel M.
APPLICANT: Mattson, Jeanine
APPLICANT: Moshrefi, Mehrdad
APPLICANT: Parham, Christi
TITLE OF INVENTION: MAMMALIAN GENES; RELATED REAGENTS
FILE REFERENCE: DX01343X
CURRENT APPLICATION NUMBER: US/10/236,055A
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/317,988
PRIOR FILING DATE: 2001-09-06
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 420
TYPE: PRT
ORGANISM: Homo sapiens
US-10-236-055A-2

Query Match 13.3%; Score 230; DB 12; Length 420;
Best Local Similarity 38.4%; Pred. No. 3.7e-13;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;
QY 163 DAPARGP-----PAPPEPSAFASTRSLVSGDAGPGRHOPLAFDTEFVNIGG 212
DB 267 DVPVTNPAATILPVHYVPLPQOMRVAFAAART-----SNLAPGTLDPVFDLLNNIGE 321
QY 213 DFDAAAGVFCRLPGAYFFSFTLGKLPR-KRTLSVKLMKNRDEVOAMITYDDGASRRREMOS 271
DB 322 TFDLQGRFNCVPVNGTYVFIHMLKLA VNPPLYVNLKNEEVLVSAYANDGAP-DHETAS 380
QY 272 QSVMLALRGGDAVWLISHDHDGYGA-YSNHGKYITFSGFLVYPD 314
DB 381 NHAITQLFGQDQIWLRLH---RGAIYSSWKXSTFSGYLLYQD 420

RESULT 7
US-09-738-973-185
Sequence 185, Application US/09738973
Patent No. US20020110563A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Henderson, Robert A.
APPLICANT: Lodes, Michael J.
APPLICANT: Flins, Steven P.
APPLICANT: Mohamath, Raodoh
APPLICANT: Algate, Paul A.
APPLICANT: Secrist, Heather
APPLICANT: Indrias, Carol Yoseph
APPLICANT: Benson, Darin R.
APPLICANT: Elliott, Mark
APPLICANT: Mannion, Jane
APPLICANT: Kaloje, Michael D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR

```
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 185
; LENGTH: 746
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-738-973-185

Query Match          13.3%; Score 230; DB 10; length 746;
Best Local Similarity 38.4%; Pred. No. 8e-13;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

QY 163 DAPARGP-----PAPPEPSAFSAARTRSLVGS D A G P G P R H O P L A F D T E F V N I G G 212
Db 593 DVPVTNPAATILPVHVYPLPQOMRVA F S A A R T-----S N L A P G T L D Q P I V F D L L N N L G E 647

QY 213 DFDAAAGVFRCLPGAYFFSFTLGKLP-RKTL SVKLMKNRDEVQAM IYDDGASRRREMOS 271
Db 648 TFDLQLGRFNCVPNGTYVFIFHMLKLA VNVPLVYVNL MKN E E V L V S A Y A N D G A P -D H E T A S 706

QY 272 QSVMLALRRGDAVWLLSHDHDGYGA-YSNHGKYITFSGFLVYPPD 314
Db 707 NHAIIQLFGQDQIWRLH---RGA IY G S S W K Y S T F S G Y L L Y O D 746

RESULT 8
US-09-854-133-185
; Sequence 185, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 185
; LENGTH: 746
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-854-133-185

Query Match          13.3%; Score 230; DB 10; length 746;
Best Local Similarity 38.4%; Pred. No. 8e-13;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

QY 163 DAPARGP-----PAPPEPSAFSAARTRSLVGS D A G P G P R H O P L A F D T E F V N I G G 212
Db 593 DVPVTNPAATILPVHVYPLPQOMRVA F S A A R T-----S N L A P G T L D Q P I V F D L L N N L G E 647

QY 213 DFDAAAGVFRCLPGAYFFSFTLGKLP-RKTL SVKLMKNRDEVQAM IYDDGASRRREMOS 271
Db 648 TFDLQLGRFNCVPNGTYVFIFHMLKLA VNVPLVYVNL MKN E E V L V S A Y A N D G A P -D H E T A S 706

QY 272 QSVMLALRRGDAVWLLSHDHDGYGA-YSNHGKYITFSGFLVYPPD 314
Db 707 NHAIIQLFGQDQIWRLH---RGA IY G S S W K Y S T F S G Y L L Y O D 746

RESULT 9
US-10-144-649A-185
; Sequence 185, Application US/10144649A
; Publication No. US20030118599A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Algate, Paul A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C11
; CURRENT APPLICATION NUMBER: US/10/144,649A
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 749
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 185
; LENGTH: 746
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-144-649A-185

Query Match          13.3%; Score 230; DB 15; length 746;
Best Local Similarity 38.4%; Pred. No. 8e-13;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

QY 163 DAPARGP-----PAPPEPSAFSAARTRSLVGS D A G P G P R H O P L A F D T E F V N I G G 212
Db 593 DVPVTNPAATILPVHVYPLPQOMRVA F S A A R T-----S N L A P G T L D Q P I V F D L L N N L G E 647

QY 213 DFDAAAGVFRCLPGAYFFSFTLGKLP-RKTL SVKLMKNRDEVQAM IYDDGASRRREMOS 271
Db 648 TFDLQLGRFNCVPNGTYVFIFHMLKLA VNVPLVYVNL MKN E E V L V S A Y A N D G A P -D H E T A S 706

QY 272 QSVMLALRRGDAVWLLSHDHDGYGA-YSNHGKYITFSGFLVYPPD 314
Db 707 NHAIIQLFGQDQIWRLH---RGA IY G S S W K Y S T F S G Y L L Y O D 746

RESULT 10
US-10-309-422-10
; Sequence 10, Application US/10309422
; Publication No. US20030139587A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: NO. US20030139587A1e1 Human Proteins and Polynucleotides Encodin
; FILE REFERENCE: LEX-0142-USA
; CURRENT APPLICATION NUMBER: US/10/309,422
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: US/09/798,771
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,557
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 908
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-309-422-10

Query Match          13.3%; Score 230; DB 12; length 908;
Best Local Similarity 38.4%; Pred. No. 1e-12;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

QY 163 DAPARGP-----PAPPEPSAFSAARTRSLVGS D A G P G P R H O P L A F D T E F V N I G G 212
Db 755 DVPVTNPAATILPVHVYPLPQOMRVA F S A A R T-----S N L A P G T L D Q P I V F D L L N N L G E 809

QY 213 DFDAAAGVFRCLPGAYFFSFTLGKLP-RKTL SVKLMKNRDEVQAM IYDDGASRRREMOS 271
Db 810 TFDLQLGRFNCVPNGTYVFIFHMLKLA VNVPLVYVNL MKN E E V L V S A Y A N D G A P -D H E T A S 868

QY 272 QSVMLALRRGDAVWLLSHDHDGYGA-YSNHGKYITFSGFLVYPPD 314
```


Db 869 NHAIIQLFGDQIWLRLH----RGAITYGSSWKYSTFSGYLLYQD 908

RESULT 11

US-10-309-422-22
; Sequence 22, Application US/10309422
; Publication No. US20030139587A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20030139587A1e1 Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: LEX-0142-USA
; CURRENT APPLICATION NUMBER: US/10/309,422
; PRIOR FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: US/09/798,771
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,557
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 909
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-309-422-22

Query Match 13.3%; Score 230; DB 12; Length 909;
Best Local Similarity 38.4%; Pred. No. 1e-12;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

QY 163 DAPARGP-----PAPPEPSAFSAARTSLVGSADGPRRHQPLAFDTEFVNIG 212
Db 756 DVPVTNPAATILPVHVYPLPQOMRVAFSAART-----SNLAPGTLDPQPIVFDLLNNLGE 810
QY 213 DFDAAGVFRCLPGAYFFSFTLGKLP-RKTLGVKLMKNRDEVQAMITYDDGASRRREMOS 271
Db 811 TFDLQGRFNCVPVNGTYVFIHMLKLA VNPVLYNLMKNEEVLVSAYANDGAP-DHETAS 869
QY 272 QSVMLALRRGDVAVWLSSHHDGYGA-YSNHGKYITFSGFLVYPD 314
Db 870 NHAIIQLFGDQIWLRLH----RGAITYGSSWKYSTFSGYLLYQD 909

RESULT 12

US-10-309-422-14
; Sequence 14, Application US/10309422
; Publication No. US20030139587A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20030139587A1e1 Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: LEX-0142-USA
; CURRENT APPLICATION NUMBER: US/10/309,422
; PRIOR FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: US/09/798,771
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,557
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 957
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-309-422-14

Query Match 13.3%; Score 230; DB 12; Length 957;
Best Local Similarity 38.4%; Pred. No. 1.1e-12;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

QY 163 DAPARGP-----PAPPEPSAFSAARTSLVGSADGPRRHQPLAFDTEFVNIG 212
Db 804 DVPVTNPAATILPVHVYPLPQOMRVAFSAART-----SNLAPGTLDPQPIVFDLLNNLGE 858
QY 213 DFDAAGVFRCLPGAYFFSFTLGKLP-RKTLGVKLMKNRDEVQAMITYDDGASRRREMOS 271
Db 859 TFDLQGRFNCVPVNGTYVFIHMLKLA VNPVLYNLMKNEEVLVSAYANDGAP-DHETAS 917
QY 272 QSVMLALRRGDVAVWLSSHHDGYGA-YSNHGKYITFSGFLVYPD 314
Db 918 NHAIIQLFGDQIWLRLH----RGAITYGSSWKYSTFSGYLLYQD 957

RESULT 13

US-10-309-422-26
; Sequence 26, Application US/10309422
; Publication No. US20030139587A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20030139587A1e1 Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: LEX-0142-USA
; CURRENT APPLICATION NUMBER: US/10/309,422
; PRIOR FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: US/09/798,771
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,557
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 958
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-309-422-26

Query Match 13.3%; Score 230; DB 12; Length 958;
Best Local Similarity 38.4%; Pred. No. 1.1e-12;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

QY 163 DAPARGP-----PAPPEPSAFSAARTSLVGSADGPRRHQPLAFDTEFVNIG 212
Db 805 DVPVTNPAATILPVHVYPLPQOMRVAFSAART-----SNLAPGTLDPQPIVFDLLNNLGE 859
QY 213 DFDAAGVFRCLPGAYFFSFTLGKLP-RKTLGVKLMKNRDEVQAMITYDDGASRRREMOS 271
Db 860 TFDLQGRFNCVPVNGTYVFIHMLKLA VNPVLYNLMKNEEVLVSAYANDGAP-DHETAS 918
QY 272 QSVMLALRRGDVAVWLSSHHDGYGA-YSNHGKYITFSGFLVYPD 314
Db 919 NHAIIQLFGDQIWLRLH----RGAITYGSSWKYSTFSGYLLYQD 958

RESULT 14

US-10-309-422-18
; Sequence 18, Application US/10309422
; Publication No. US20030139587A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20030139587A1e1 Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: LEX-0142-USA
; CURRENT APPLICATION NUMBER: US/10/309,422
; PRIOR FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: US/09/798,771
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,557
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18

LENGTH: 992
TYPE: PRT
ORGANISM: homo sapiens
US-10-309-422-18

Query Match 13.3%; Score 230; DB 12; Length 992;
Best Local Similarity 38.4%; Pred. No. 1.2e-12;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

QY 163 DAPARGP-----PAPPEPSAFSAARTRSLVSGDAGGPRHQPLAFDTEFVNIGG 212
| | | | | : | | | | | : | | | | | : | | | | | : | |
Db 839 DVPVTNPATILPVHVYPLPQGMVAFSAART-----SNLAPGTLDDQPIVFDLLNNLGE 893
| | | | | : | | | | | : | | | | | : | | | | | : | |
QY 213 DFDAAAGVFCRLPGAYFFSFTLGKLP-RKTL SVKLMKNRDEVQAMTYDDGASRRREMOS 271
| | | | | : | | | | | : | | | | | : | | | | | : | |
Db 894 TFDLQLGRFNCVPVNGTYVFIFHMLKLA VNVPLYVNL MKNEEVLVSAYANDGAP-DHETAS 952
| | | | | : | | | | | : | | | | | : | | | | | : | |
QY 272 QSVMLALRGGDAVWLLSHDHGCGA-YSNHGKYITFSGFLVYPD 314
: | | | | : | | | | : | | | | : | | | | : | | | | : | | | |
Db 953 NHAIIQLFQGDQIWLRH---RGAIIYGSSWKYSTFSGYLLYQD 992

RESULT 15
US-10-309-422-30
; Sequence 30, Application US/10309422
; Publication No. US20030139587A1
; GENERAL INFORMATION:
; APPLICANT: walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20030139587A1 Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: LEX-0142-USA
; CURRENT APPLICATION NUMBER: US/10/309,422
; PRIOR FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: US/09/798,771
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,557
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 30
; LENGTH: 993
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-309-422-30

Query Match 13.3%; Score 230; DB 12; Length 993;
Best Local Similarity 38.4%; Pred. No. 1.2e-12;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

QY 163 DAPARGP-----PAPPEPSAFSAARTRSLVSGDAGGPRHQPLAFDTEFVNIGG 212
| | | | | : | | | | | : | | | | | : | | | | | : | |
Db 840 DVPVTNPATILPVHVYPLPQGMVAFSAART-----SNLAPGTLDDQPIVFDLLNNLGE 894
| | | | | : | | | | | : | | | | | : | | | | | : | |
QY 213 DFDAAAGVFCRLPGAYFFSFTLGKLP-RKTL SVKLMKNRDEVQAMTYDDGASRRREMOS 271
| | | | | : | | | | | : | | | | | : | | | | | : | |
Db 895 TFDLQLGRFNCVPVNGTYVFIFHMLKLA VNVPLYVNL MKNEEVLVSAYANDGAP-DHETAS 953
| | | | | : | | | | | : | | | | | : | | | | | : | |
QY 272 QSVMLALRGGDAVWLLSHDHGCGA-YSNHGKYITFSGFLVYPD 314
: | | | | : | | | | : | | | | : | | | | : | | | | : | | | |
Db 954 NHAIIQLFQGDQIWLRH---RGAIIYGSSWKYSTFSGYLLYQD 993

Search completed: January 12, 2004, 08:33:35
Job time : 42.0579 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:14 ; Search time 16.7455 Seconds
(without alignments)
831.284 Million cell updates/sec

Title: US-10-085-167-2

Perfect score: 1726

Sequence: 1 MLPLLGLLGPACWALGPT.....LVYPDLAPAAPGLGASELL 329

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---|
| 1 | 230 | 13.3 | 746 | 4 | US-09-370-838-185 Sequence 185, App |
| 2 | 220.5 | 12.8 | 281 | 3 | US-09-118-408-44 Sequence 44, Appl |
| 3 | 220.5 | 12.8 | 281 | 4 | US-09-506-855-44 Sequence 44, Appl |
| 4 | 220.5 | 12.8 | 281 | 4 | US-09-911-176B-44 Sequence 44, Appl |
| 5 | 220.5 | 12.8 | 281 | 4 | US-09-619-740-44 Sequence 44, Appl |
| 6 | 220.5 | 12.8 | 281 | 4 | US-09-506-852-44 Sequence 44, Appl |
| 7 | 215 | 12.5 | 281 | 3 | US-09-118-408-2 Sequence 2, Appli |
| 8 | 215 | 12.5 | 281 | 4 | US-09-506-855-2 Sequence 2, Appli |
| 9 | 215 | 12.5 | 281 | 4 | US-09-911-176B-2 Sequence 2, Appli |
| 10 | 215 | 12.5 | 281 | 4 | US-09-619-740-2 Sequence 2, Appli |
| 11 | 215 | 12.5 | 281 | 4 | US-09-506-852-2 Sequence 2, Appli |
| 12 | 213 | 12.3 | 228 | 4 | US-09-336-536-4 Sequence 4, Appli |
| 13 | 213 | 12.3 | 243 | 3 | US-09-140-804-2 Sequence 2, Appli |
| 14 | 213 | 12.3 | 243 | 4 | US-09-336-536-3 Sequence 3, Appli |
| 15 | 213 | 12.3 | 243 | 4 | US-09-686-838B-2 Sequence 2, Appli |
| 16 | 207 | 12.0 | 243 | 3 | US-09-188-930-295 Sequence 295, App |
| 17 | 207 | 12.0 | 243 | 4 | US-09-312-283C-295 Sequence 295, App |
| 18 | 207 | 12.0 | 294 | 3 | US-09-188-930-294 Sequence 294, App |
| 19 | 207 | 12.0 | 294 | 4 | US-09-312-283C-294 Sequence 294, App |
| 20 | 204 | 11.8 | 285 | 4 | US-09-312-283C-382 Sequence 382, App |
| 21 | 202 | 11.7 | 228 | 4 | US-09-336-536-11 Sequence 11, Appl |
| 22 | 202 | 11.7 | 243 | 4 | US-09-336-536-10 Sequence 10, Appl |
| 23 | 200 | 11.6 | 259 | 4 | US-09-996-243-47 Sequence 47, Appl |
| 24 | 198 | 11.5 | 260 | 4 | US-09-489-847-198 Sequence 198, App |
| 25 | 198 | 11.5 | 287 | 4 | US-09-489-847-349 Sequence 349, App |
| 26 | 196 | 11.4 | 231 | 4 | US-09-530-423-2 Sequence 2, Appli |
| 27 | 196 | 11.4 | 244 | 2 | US-08-463-911-7 Sequence 7, Appli |

| | | | | | | |
|----|-----|------|-----|---|-------------------|-------------------|
| 28 | 196 | 11.4 | 244 | 3 | US-09-140-804-3 | Sequence 3, Appli |
| 29 | 196 | 11.4 | 244 | 4 | US-09-336-536-20 | Sequence 20, Appl |
| 30 | 196 | 11.4 | 244 | 4 | US-09-530-423-1 | Sequence 1, Appli |
| 31 | 196 | 11.4 | 244 | 4 | US-09-686-838B-3 | Sequence 3, Appli |
| 32 | 196 | 11.4 | 244 | 4 | US-09-911-176B-48 | Sequence 48, Appl |
| 33 | 196 | 11.4 | 244 | 4 | US-09-552-225A-3 | Sequence 3, Appli |
| 34 | 196 | 11.4 | 244 | 4 | US-09-619-740-51 | Sequence 51, Appl |
| 35 | 196 | 11.4 | 244 | 4 | US-09-776-976-6 | Sequence 6, Appli |
| 36 | 196 | 11.4 | 244 | 4 | US-09-909-547-6 | Sequence 6, Appli |
| 37 | 196 | 11.4 | 244 | 4 | US-09-569-852B-6 | Sequence 6, Appli |
| 38 | 196 | 11.4 | 247 | 2 | US-08-463-911-2 | Sequence 2, Appli |
| 39 | 196 | 11.4 | 247 | 3 | US-09-140-804-8 | Sequence 8, Appli |
| 40 | 196 | 11.4 | 247 | 3 | US-09-118-408-3 | Sequence 3, Appli |
| 41 | 196 | 11.4 | 247 | 4 | US-09-506-855-3 | Sequence 3, Appli |
| 42 | 196 | 11.4 | 247 | 4 | US-09-686-838B-8 | Sequence 8, Appli |
| 43 | 196 | 11.4 | 247 | 4 | US-09-911-176B-3 | Sequence 3, Appli |
| 44 | 196 | 11.4 | 247 | 4 | US-09-619-740-3 | Sequence 3, Appli |
| 45 | 196 | 11.4 | 247 | 4 | US-09-776-976-4 | Sequence 4, Appli |

ALIGNMENTS

RESULT 1
US-09-370-838-185
; Sequence 185, Application US/09370838
; Patent No. 644425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370, 838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285, 323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 185
; LENGTH: 746
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-370-838-185

Query Match 13.3%; Score 230, DB 4, Length 746;
Best Local Similarity 38.4%; Pred. No. 9.1e-15;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

QY 163 DAPARGP-----PAPPEPRSAFSAARTSLVGS DAGCPRHQPLAFDTEFNIGG 212
Db DVPVTNPATITLVHVYPLPQOMRVAFSAART-----SNLAPGTLDPYFDLLNLIGE 647

QY 213 DFDAAGVFRCHLPGAYFFSFTLGKLP-RKTLVYKLMKNRDEVQAMTYDGDASRRREMOS 271
Db TFDLQLGRFNCPTNGTYVFIFHMLKLA VNVPLVYVNLKNEEVLVSAYANDGAP-DHETAS 706

QY 272 QSVMLARRGDVAWLLSHDHDGYGA-YSNHKXITFSGFLVYPP 314
Db NHAILQLFGQDQIWRLH----RGAITYSSWKSTFSGYLLYQD 746

RESULT 2
US-09-118-408-44
; Sequence 44, Application US/09118408A
; Patent No. 6265544
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
; FILE REFERENCE: 97-30
; CURRENT APPLICATION NUMBER: US/09/118, 408A

Db 259 IFSDEFDTYITFSGYLVPASEP 281

RESULT 6

US-09-506-852-44

; Sequence 44, Application US/09506852

; Patent No. 6566499

; GENERAL INFORMATION:

; APPLICANT: Sheppard, Paul O.

; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS

; FILE REFERENCE: 97-30

; CURRENT APPLICATION NUMBER: US/09/506,852

; CURRENT FILING DATE: 2000-02-17

; EARLIER APPLICATION NUMBER: 60/053,154

; EARLIER FILING DATE: 1997-07-18

; NUMBER OF SEQ ID NOS: 44

; SOFTWARE: FastSeq for windows Version 3.0

; SEQ ID NO 44

; LENGTH: 281

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-506-852-44

Query Match 12.8%; Score 220.5; DB 4; Length 281;

Best Local Similarity 39.2%; Pred. No. 2.1e-14;

Matches 56; Conservative 20; Mismatches 58; Indels 9; Gaps 4;

QY 177 SAFSARTSLVGSADGPGPRHQPLAFDTEFVNIGDFAAGVFRCLPGAYFFSFTLG 236

Db 146 AAFSVGRKKALHSND-----YFQPVFVDTFVNLVKHFMFTGKFCYCVPGIYFSLNVH 200

QY 237 KLPRKTLVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRGDAVW--ILSHDHG 294

Db 201 TWNOKETYLHIMKNEEV-VILYAQ-VSDRSIMOSQSLMELREDEVWVRLFKGERENA 258

QY 295 GAYSNHGKYITFSGFLVYPDLP 317

Db 259 IFSDEFDTYITFSGYLVPASEP 281

RESULT 7

US-09-118-408-2

; Sequence 2, Application US/09118408A

; Patent No. 6265544

; GENERAL INFORMATION:

; APPLICANT: Sheppard, Paul O.

; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS

; FILE REFERENCE: 97-30

; CURRENT APPLICATION NUMBER: US/09/118,408A

; CURRENT FILING DATE: 1998-07-17

; EARLIER APPLICATION NUMBER: 60/053,154

; EARLIER FILING DATE: 1997-07-18

; NUMBER OF SEQ ID NOS: 47

; SOFTWARE: FastSeq for windows Version 3.0

; SEQ ID NO 2

; LENGTH: 281

; TYPE: PRT

; ORGANISM: Homo sapien

US-09-118-408-2

Query Match 12.5%; Score 215; DB 3; Length 281;

Best Local Similarity 30.2%; Pred. No. 7.6e-14;

Matches 77; Conservative 26; Mismatches 86; Indels 66; Gaps 11;

QY 69 RCRVPGAYFFSFTAGKAPHSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDY 128

Db 75 RCCDPGTSMPATA--VPQINITILKGEKGR-----GDRG-----LQGY 113

QY 129 GDT--VWLRLHGAP--HYALGAPGATFSGYLVAADADADAPARGPAPPEPRSAFSAAR 183

Db 114 GKTGSAGARGHTGPKGQKSGMGA GERCKSH--YA-----AFSVGR 152

QY 184 TRSLVGSADGPGPRH-----QPLAFDTEFVNIGDFAAGVFRCLPGAYFFSFTLGKL 238

Db 153 KK-----PMHSNHYQTIVIFDTEFVNLVDHFNMTGKFCYCVPGIYFSLNVHTW 202

QY 239 PRKTLVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRGDAVWILSHDHG 298

Db 203 NOKETYLHIMKNEEVILFAQVG--DRSIMOSQSLMELREQDQVWVRLYKGERENAIF 260

QY 299 NH--GKYITFSGFLV 311

Db 261 SEELDTYITFSGYL 275

RESULT 8

US-09-506-855-2

; Sequence 2, Application US/09506855

; Patent No. 6448221

; GENERAL INFORMATION:

; APPLICANT: Sheppard, Paul O.

; APPLICANT: Lasser, Gerald W.

; APPLICANT: Bishop, Paul D.

; TITLE OF INVENTION: INHIBITORS FOR USE IN HEMOSTASIS AND

; FILE REFERENCE: 99-12

; CURRENT APPLICATION NUMBER: US/09/506,855

; CURRENT FILING DATE: 2000-02-17

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: FastSeq for windows Version 3.0

; SEQ ID NO 2

; LENGTH: 281

; TYPE: PRT

; ORGANISM: Homo sapien

US-09-506-855-2

Query Match 12.5%; Score 215; DB 4; Length 281;

Best Local Similarity 30.2%; Pred. No. 7.6e-14;

Matches 77; Conservative 26; Mismatches 86; Indels 66; Gaps 11;

QY 69 RCRVPGAYFFSFTAGKAPHSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDY 128

Db 75 RCCDPGTSMPATA--VPQINITILKGEKGR-----GDRG-----LQGY 113

QY 129 GDT--VWLRLHGAP--HYALGAPGATFSGYLVAADADADAPARGPAPPEPRSAFSAAR 183

Db 114 GKTGSAGARGHTGPKGQKSGMGA GERCKSH--YA-----AFSVGR 152

QY 184 TRSLVGSADGPGPRH-----QPLAFDTEFVNIGDFAAGVFRCLPGAYFFSFTLGKL 238

Db 153 KK-----PMHSNHYQTIVIFDTEFVNLVDHFNMTGKFCYCVPGIYFSLNVHTW 202

QY 239 PRKTLVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRGDAVWILSHDHG 298

Db 203 NOKETYLHIMKNEEVILFAQVG--DRSIMOSQSLMELREQDQVWVRLYKGERENAIF 260

QY 299 NH--GKYITFSGFLV 311

Db 261 SEELDTYITFSGYL 275

RESULT 9

US-09-911-176B-2

; Sequence 2, Application US/09911176B

; Patent No. 6518403

; GENERAL INFORMATION:

; APPLICANT: Sheppard, Paul O.

; TITLE OF INVENTION: ANTIBODIES THAT BIND AN

; FILE REFERENCE: 97-30D1

; CURRENT APPLICATION NUMBER: US/09/911,176B

; CURRENT FILING DATE: 2001-07-23

; PRIOR APPLICATION NUMBER: 09/118,408

; PRIOR FILING DATE: 1998-07-17

; PRIOR APPLICATION NUMBER: 60/053,154

; PRIOR FILING DATE: 1997-07-18

```
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-911-176B-2
```

```
Query Match          12.5%; Score 215; DB 4; Length 281;
Best Local Similarity 30.2%; Pred. No. 7.6e-14;
Matches 77; Conservative 26; Mismatches 86; Indels 66; Gaps 11;
```

```
QY 69 RCRVPGAYFFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDY 128
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 75 RCCDPGTSMPATA--VPQINITILKEKGD-----GDRG-----LQCKY 113

QY 129 GDT--VWLRHGA---HYALGAPGATFSGYLVYADADADAPARGPAPPEPRSAFSAAR 183
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 114 GKTGSAGARGHTGPKGQKSGMGAPGERCKSH--YA-----AFSVGR 152

QY 184 TRSLVGS DAGPGRH-----QPLAFDTEFVNIGGDFDAAGVFRCLPGAYFFSFTLGKL 238
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 153 KK-----PMHSNHYYQTIVFDTEFVNLYDHFNMTGKFCYVPGLYFFSLNVHTW 202

QY 239 PRKTLVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRRGDVWVLSHDHDCGYAYS 298
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 203 NQKETYLHIMKNEEVLILFAQVG--DRSIMQSQSLMLELRQDQVWVRLYKGERENAIF 260

QY 299 NH--GKYITFSGFLV 311
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 261 SEELDTYITFSGYLV 275
```

RESULT 10

```
US-09-619-740-2
; Sequence 2, Application US/09619740
; Patent No. 6544946
; GENERAL INFORMATION:
; APPLICANT: Shepard, Paul O.
; APPLICANT: Laesser, Gerald W.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: INHIBITORS FOR USE IN HEMOSTASIS AND IMMUNE FUNCTION
; FILE REFERENCE: 99-12C3
; CURRENT APPLICATION NUMBER: US/09/619,740
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/253,604
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/444,794
; PRIOR FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: 09/506,855
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-619-740-2
```

```
Query Match          12.5%; Score 215; DB 4; Length 281;
Best Local Similarity 30.2%; Pred. No. 7.6e-14;
Matches 77; Conservative 26; Mismatches 86; Indels 66; Gaps 11;
```

```
QY 69 RCRVPGAYFFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDY 128
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 75 RCCDPGTSMPATA--VPQINITILKEKGD-----GDRG-----LQCKY 113

QY 129 GDT--VWLRHGA---HYALGAPGATFSGYLVYADADADAPARGPAPPEPRSAFSAAR 183
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 114 GKTGSAGARGHTGPKGQKSGMGAPGERCKSH--YA-----AFSVGR 152

QY 184 TRSLVGS DAGPGRH-----QPLAFDTEFVNIGGDFDAAGVFRCLPGAYFFSFTLGKL 238
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
Db 153 KK-----PMHSNHYYQTIVFDTEFVNLYDHFNMTGKFCYVPGLYFFSLNVHTW 202

QY 239 PRKTLVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRRGDVWVLSHDHDCGYAYS 298
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 203 NQKETYLHIMKNEEVLILFAQVG--DRSIMQSQSLMLELRQDQVWVRLYKGERENAIF 260

QY 299 NH--GKYITFSGFLV 311
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 261 SEELDTYITFSGYLV 275
```

RESULT 11

```
US-09-506-852-2
; Sequence 2, Application US/09506852
; Patent No. 6566499
; GENERAL INFORMATION:
; APPLICANT: Shepard, Paul O.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
; FILE REFERENCE: 97-30
; CURRENT APPLICATION NUMBER: US/09/506,852
; PRIOR FILING DATE: 2000-02-17
; EARLIER APPLICATION NUMBER: 60/053,154
; EARLIER FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-506-852-2
```

```
Query Match          12.5%; Score 215; DB 4; Length 281;
Best Local Similarity 30.2%; Pred. No. 7.6e-14;
Matches 77; Conservative 26; Mismatches 86; Indels 66; Gaps 11;
```

```
QY 69 RCRVPGAYFFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDY 128
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 75 RCCDPGTSMPATA--VPQINITILKEKGD-----GDRG-----LQCKY 113

QY 129 GDT--VWLRHGA---HYALGAPGATFSGYLVYADADADAPARGPAPPEPRSAFSAAR 183
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 114 GKTGSAGARGHTGPKGQKSGMGAPGERCKSH--YA-----AFSVGR 152

QY 184 TRSLVGS DAGPGRH-----QPLAFDTEFVNIGGDFDAAGVFRCLPGAYFFSFTLGKL 238
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 153 KK-----PMHSNHYYQTIVFDTEFVNLYDHFNMTGKFCYVPGLYFFSLNVHTW 202

QY 239 PRKTLVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRRGDVWVLSHDHDCGYAYS 298
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 203 NQKETYLHIMKNEEVLILFAQVG--DRSIMQSQSLMLELRQDQVWVRLYKGERENAIF 260

QY 299 NH--GKYITFSGFLV 311
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 261 SEELDTYITFSGYLV 275
```

RESULT 12

```
US-09-336-536-4
; Sequence 4, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 228
; TYPE: PRT
```


| | | | |
|----|-----|-----------|-----|
| Qy | 306 | FSGFLVYPD | 314 |
| | | | |
| Db | 227 | FSGFLVYSD | 235 |

Search completed: January 12, 2004, 08:20:03
Job time : 17.7455 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:09 ; Search time 0.926148 Seconds
(without alignment)
1661.397 Million cell updates/sec

Title: US-10-085-167-2_COPY_1_16
Perfect score: 87
Sequence: 1 MLPLLLGLGPACWA 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|--------|--------------------|
| 1 | 57 | 65.5 | 993 | 2 | I48653 | mouse developmenta |
| 2 | 52 | 59.8 | 115 | 2 | S72634 | type 4 fimbrial bi |
| 3 | 49 | 56.3 | 1180 | 2 | T31066 | vascular cadherin- |
| 4 | 46 | 52.9 | 283 | 2 | T16945 | hypothetical prote |
| 5 | 45 | 51.7 | 303 | 2 | AE3201 | conserved hypothet |
| 6 | 45 | 51.7 | 315 | 2 | E75382 | hypothetical prote |
| 7 | 45 | 51.7 | 416 | 1 | S06763 | calreticulin precu |
| 8 | 45 | 51.7 | 416 | 1 | JH0819 | calreticulin precu |
| 9 | 45 | 51.7 | 418 | 1 | A34154 | calreticulin precu |
| 10 | 45 | 51.7 | 628 | 2 | D97073 | PTS system, beta-g |
| 11 | 44 | 50.6 | 318 | 2 | A55429 | 11-cis retinol deh |
| 12 | 44 | 50.6 | 553 | 2 | F75407 | probable transport |
| 13 | 44 | 50.6 | 1220 | 2 | T18291 | patched protein - |
| 14 | 43 | 49.4 | 119 | 2 | C81919 | probable membrane |
| 15 | 43 | 49.4 | 119 | 2 | F81182 | conserved hypothet |
| 16 | 43 | 49.4 | 146 | 2 | B75021 | hypothetical prote |
| 17 | 43 | 49.4 | 159 | 2 | G75555 | conserved hypothet |
| 18 | 43 | 49.4 | 273 | 1 | C70127 | guFA protein homol |
| 19 | 43 | 49.4 | 275 | 2 | H83325 | hypothetical prote |
| 20 | 43 | 49.4 | 341 | 1 | VVVPK2 | coat protein VP2 - |
| 21 | 43 | 49.4 | 344 | 1 | VVVP2J | coat protein VP2 - |
| 22 | 43 | 49.4 | 351 | 1 | VVVP2B | coat protein VP2 - |
| 23 | 43 | 49.4 | 351 | 1 | VVVPAS | coat protein VP2 - |
| 24 | 43 | 49.4 | 352 | 1 | VVVP24 | coat protein VP2 - |
| 25 | 43 | 49.4 | 368 | 2 | S75652 | ABC-type transport |
| 26 | 43 | 49.4 | 461 | 2 | T11829 | NADH2 dehydrogenas |
| 27 | 43 | 49.4 | 558 | 2 | AG2626 | ABC transporter, m |
| 28 | 43 | 49.4 | 558 | 2 | F97408 | hypothetical prote |
| 29 | 42.5 | 48.9 | 319 | 1 | VVVP2 | coat protein VP2 - |

| | | | | | | |
|----|------|------|-----|---|--------|--------------------|
| 30 | 42.5 | 48.9 | 319 | 1 | VVVP2 | coat protein VP2 - |
| 31 | 42 | 48.3 | 227 | 2 | JQ1997 | hypothetical 26k p |
| 32 | 42 | 48.3 | 361 | 2 | AE1979 | ABC transporter Ar |
| 33 | 42 | 48.3 | 403 | 2 | T34923 | hypothetical ROK f |
| 34 | 42 | 48.3 | 657 | 2 | D71351 | probable primosoma |
| 35 | 41.5 | 47.7 | 265 | 2 | B96913 | ABC-type probable |
| 36 | 41 | 47.1 | 66 | 2 | A83575 | hypothetical prote |
| 37 | 41 | 47.1 | 110 | 2 | S22893 | T-cell receptor al |
| 38 | 41 | 47.1 | 121 | 2 | H81892 | hypothetical prote |
| 39 | 41 | 47.1 | 229 | 2 | T18629 | phosphatidate cyti |
| 40 | 41 | 47.1 | 265 | 2 | A82000 | phosphatidate cyti |
| 41 | 41 | 47.1 | 265 | 2 | B81229 | phosphatidate cyti |
| 42 | 41 | 47.1 | 267 | 2 | T36675 | probable integral |
| 43 | 41 | 47.1 | 277 | 2 | I37552 | OX40 homolog - hum |
| 44 | 41 | 47.1 | 341 | 2 | A69209 | conserved hypothet |
| 45 | 41 | 47.1 | 345 | 1 | VVVP2H | coat protein VP2 - |

ALIGNMENTS

RESULT 1
I48653
mouse developmental kinase 5 (MDK5) - mouse
C/Species: Mus musculus (house mouse)
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999
C/Accession: I48653
R/Closetek, T.; Lerch, M.M.; Ullrich, A.
Oncogene 11, 2085-2095, 1995
A/Title: Cloning, characterization, and differential expression of MDK2 and MDK5, two nov
A/Reference number: I48652; MUID:96074837; PMID:7478528
A/Accession: I48653
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-993 <RES>
A/Cross-references: EMBL:Z49086; NID:gl089899; PIDN:CAA88910.1; PID:gl089900
C/Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat hc
C/Keywords: ATP; transmembrane protein
F/626-894/Domain: protein kinase homology <KIN>
F/634-642/Region: protein kinase ATP-binding motif
F/917-983/Domain: SAM homology <SAM>

Query Match 65.5%; Score 57; DB 2; Length 993;
Best Local Similarity 68.8%; Pred. No. 0.73;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MLPLLLGLGPACWA 16
: ||||| |||||
Db 14 LAPLLPLLLPACWA 29

RESULT 2
S72634
type 4 fimbrial biogenesis protein PilY2 PA4555 [imported] - Pseudomonas aeruginosa
C/Species: Pseudomonas aeruginosa
C/Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 31-Dec-2000
C/Accession: S72634; B83076
R/Alm, R.A.; Hallinan, J.P.; Watson, A.A.; Mattick, J.S.
Mol. Microbiol. 22, 161-173, 1996
A/Title: Fimbrial biogenesis genes of Pseudomonas aeruginosa: pilW and pilX increase the
ue.
A/Reference number: S72631; MUID:97055431; PMID:8899718
A/Accession: S72634
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-115 <ALM>
A/Cross-references: EMBL:L76605; NID:gl246298; PIDN:AAA93503.1; PID:gl246303
A/Experimental source: strain PAO1
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1996
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bri
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Latbig, K.; Lim,
.; LoY, S.; Olson, M.V.
Nature 406, 959-964, 2000

A/Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
A/Reference number: AB2950; MUID:20437337; PMID:10984043
A/Accession: E83076
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-115 <STO>
A/Cross-references: GB:AE004869; GB:AE004091; NID:g9950793; PIDN:AG07943.1; GSPDB:GN001
A/Experimental source: strain PA01
C/Genetics:
A/Gene: pilY2; PA4555
C/Superfamily: *Pseudomonas aeruginosa* pilY2 protein
C/Keywords: fimbria

Query Match 59.8%; Score 52; DB 2; Length 115;
Best Local Similarity 56.2%; Pred. No. 0.6;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MLPLLGLGPACWA 16
:|||||
Db 3 VLPMLLALAVPGLCWA 18

RESULT 3

T31066
vascular cadherin-2 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000
C/Accession: T31066
R/Telo, P.; Breviario, F.; Huber, P.; Panzeri, C.; Dejana, E.
J. Biol. Chem. 273, 17565-17572, 1998
A/Title: Identification of a novel cadherin (vascular endothelial cadherin-2) located at
A/Reference number: Z20962; MUID:98316322; PMID:9651350
A/Accession: T31066
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-1180 <TEL>
A/Cross-references: EMBL:Y08715; NID:g2764993; PIDN:CAA69965.1; PID:g2764994
A/Experimental source: brain capillary

Query Match 56.3%; Score 49; DB 2; Length 1180;
Best Local Similarity 69.2%; Pred. No. 14;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLPLLGLGPAA 13
:|||||
Db 4 LLPPLGLGPGS 16

RESULT 4

T16945
hypothetical protein T27F7.3 - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C/Accession: T16945
R/Hallworth, K.
submitted to the EMBL Data Library, May 1996
A/Description: The sequence of *C. elegans* cosmid T27F7.
A/Reference number: Z18613
A/Accession: T16945
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-283 <HAL>
A/Cross-references: EMBL:U58762; NID:g1330398; PID:g1330400; PIDN:AAB00719.1; GSPDB:GN00
A/Experimental source: strain Bristol N2; clone T27F7
C/Genetics:
A/Gene: CESP:T27F7.3
A/Map position: 2
A/Introns: 11/1; 45/1; 63/1; 130/3; 248/3

Query Match 52.9%; Score 46; DB 2; Length 283;
Best Local Similarity 63.6%; Pred. No. 11;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPLLGLGP 11
|:|:|
Db 162 MIPILVGLGP 172

RESULT 5

AE3201
conserved hypothetical protein Atus339 [imported] - *Agrobacterium tumefaciens* (strain C58)
C/Species: *Agrobacterium tumefaciens*
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C/Accession: AE3201
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McTielle
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: AE3201
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-303 <KUR>
A/Cross-references: GB:AE008687; PIDN:AAL46027.1; PID:g17743784; GSPDB:GN00188
A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Gene: Atu5339
A/Genome: plasmid

Query Match 51.7%; Score 45; DB 2; Length 303;
Best Local Similarity 81.8%; Pred. No. 17;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 LLGLGPAC 14
:|||||
Db 70 ILGLGMAAC 80

RESULT 6

E75382
hypothetical protein - *Deinococcus radiodurans* (strain R1)
C/Species: *Deinococcus radiodurans*
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000
C/Accession: E75382
R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; I
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A/Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
A/Reference number: A75250; MUID:20036896; PMID:10567266
A/Accession: E75382
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-315 <WHI>
A/Cross-references: GB:AE001998; GB:AE000513; NID:g6459302; PIDN:AAF11113.1; PID:g645931;
A/Experimental source: strain R1
C/Genetics:
A/Gene: DR1541
A/Map position: 1
C/Superfamily: *Deinococcus radiodurans* hypothetical protein DR1541

Query Match 51.7%; Score 45; DB 2; Length 315;
Best Local Similarity 53.3%; Pred. No. 18;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MLPLLGLGPACW 15
:|||||
Db 26 LVPLAPLGLVVCW.40

RESULT 7

S06763
calreticulin precursor - mouse

N:Alternate names: 55K calcium-binding reticuloplasmin; calregulin
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S06763; JCI444; PC1233; A57498
R:Smith, M.J.; Koch, G.L.E.
EMBO J. 8, 3581-3586, 1989
A:Title: Multiple zones in the sequence of calreticulin (CRP55, calregulin, HACBP), a me
A:Reference number: S06763; MUID:90059955; PMID:2583110
A:Accession: S06763
A:Molecule type: DNA
A:Residues: 1-416 <SMI>
A:Cross-references: EMBL:X14926; NID:g50567; PIDN:CAA33053.1; PID:g50568
R:Mazarella, R.A.; Gold, P.; Cunningham, M.; Green, M.
Gene 120, 217-225, 1992
A:Title: Determination of the sequence of an expressible cDNA clone encoding ERp60/calre
A:Reference number: JCI444; MUID:93013037; PMID:1398135
A:Accession: JCI444
A:Molecule type: mRNA
A:Residues: 1-416 <MAZ>
A:Cross-references: GB:M92988; NID:g193084; PIDN:AAA37569.1; PID:g193085
A:Accession: PC1233
A:Molecule type: protein
A:Residues: 18-41 <MA2>
R:White, T.K.; Zhu, Q.; Tanzer, M.L.
J. Biol. Chem. 270, 15926-15929, 1995
A:Title: Cell surface calreticulin is a putative mannoside lectin which triggers mouse m
A:Reference number: A57498; MUID:95332280; PMID:7608143
A:Accession: A57498
A:Status: preliminary
A:Molecule type: protein
A:Residues: 74-80;142-151;186-193 <WHI>
C:Superfamily: calreticulin
C:Keywords: calcium binding
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-416/Product: calregulin #status experimental <MAT>
F:413-416/Region: endoplasmic reticulum retention signal

Query Match 51.7%; Score 45; DB 1; Length 416;
Best Local Similarity 83.3%; Pred. No. 23;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPLLLGLGPAA 13
:|||||||
Db 5 VPLLGLGLGAA 16

RESULT 8
JH0819
calreticulin precursor - rat
N:Alternate names: calcium-binding protein 3
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Sep-1993 #sequence revision 20-Aug-1994 #text change 20-Jun-2000
C:Accession: JH0819; A49176; S11205; PC1109; S45036; S04867; S39372; A34473; S13045
R:Nakamura, M.; Moriya, M.; Baba, T.; Michikawa, Y.; Yamamoto, T.; Arai, K.; Okinaga, S.
Exp. Cell Res. 205, 101-110, 1993
A:Title: An endoplasmic reticulum protein, calreticulin, is transported into the acrosom
A:Reference number: A49176; MUID:93202172; PMID:8453984
A:Accession: JH0819
A:Molecule type: mRNA
A:Residues: 1-416 <NAK>
A:Cross-references: GB:D78308; NID:g1089798; PIDN:BA11345.1; PID:g1845572
A:Accession: A49176
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-416 <NA2>
A:Cross-references: GB:D78308; NID:g1089798; PIDN:BA11345.1; PID:g1845572
A:Experimental source: Sprague-Dawley, spermatogenic cells
A>Note: sequence extracted from NCBI backbone (NCBIN:127639, NCBI:P:127643)
R:Murthy, K.K.; Banville, D.; Srikanth, C.B.; Carrier, F.; Holmes, C.; Bell, A.; Patel, Y
Nucleic Acids Res. 18, 4933, 1990
A:Title: Structural homology between the rat calreticulin gene product and the Onchocerc
A:Reference number: S11205; MUID:90370496; PMID:2395661
A:Accession: S11205

A:Molecule type: mRNA
A:Residues: 1-416 <MUR>
A:Cross-references: EMBL:X53363; NID:g55854; PIDN:CAA37446.1; PID:g55855
R:Nakamura, M.; Michikawa, Y.; Baba, T.; Okinaga, S.; Arai, K.
Biochem. Biophys. Res. Commun. 186, 668-673, 1992
A:Title: Calreticulin is present in the acrosome of spermatozoa of rat testis.
A:Reference number: PC1109; MUID:92360010; PMID:1497655
A:Accession: PC1109
A:Molecule type: protein
A:Residues: 18-32 <NAK2>
A:Experimental source: testis, strain Sprague-Dawley
R:Soennichsen, B.; Fuellekrug, J.; van Nguyen, P.; Diekmann, W.; Robinson, D.G.; Mieskes,
submitted to the EMBL Data Library, May 1994
A:Description: Retention and retrieval: both mechanisms cooperate to maintain calreticul
A:Reference number: S45036
A:Accession: S45036
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-416 <SOB>
A:Cross-references: EMBL:X79327; NID:g488840; PIDN:CAA55890.1; PID:g488841
R:Lone, Y.C.; Bailly, A.; Latruffe, N.
submitted to the EMBL Data Library, December 1988
A:Reference number: S04867
A:Accession: S04867
A:Molecule type: mRNA
A:Residues: 'R', 270-358, 'AAG' <LON>
A:Cross-references: EMBL:X13702; NID:g56055; PIDN:CAA31987.1; PID:g930260
A>Note: the authors designated the protein as D-beta-hydroxybutyrate dehydrogenase
R:Yokoi, T.; Nagayama, S.; Kajiwara, R.; Kawaguchi, Y.; Horiochi, R.; Kametaki, T.
Biochim. Biophys. Acta 1158, 339-344, 1993
A:Title: Identification of protein disulfide isomerase and calreticulin as autoimmune ant
A:Reference number: S39371; MUID:94072621; PMID:8251535
A:Accession: S39372
A:Molecule type: protein
A:Residues: 18-23, 'X', 25-32 <YOK>
R:Van, P.N.; Peter, F.; Soeling, H.D.
J. Biol. Chem. 264, 17494-17501, 1989
A:Title: Four intracisternal calcium-binding glycoproteins from rat liver microsomes with
ative calcium sequestering rat liver vesicles.
A:Reference number: A34473; MUID:9008920; PMID:2793869
A:Accession: A34473
A:Status: preliminary
A:Molecule type: protein
A:Residues: 18-36 <VAN>
R:Treves, S.; de Mattei, M.; Lanfredi, M.; Villa, A.; Green, N.M.; MacLennan, D.H.; Meld
Biochem. J. 271, 473-480, 1990
A:Title: Calreticulin is a candidate for a calsequestrin-like function in Ca(2+)-storage
A:Reference number: S13045; MUID:91054414; PMID:2241926
A:Accession: S13045
A:Molecule type: protein
A:Residues: 18-29 <TRB>
C:Superfamily: calreticulin
C:Keywords: calcium binding; glycoprotein
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-416/Product: calreticulin #status experimental <MAT>
F:204-212/Region: nuclear location signal
F:413-416/Region: endoplasmic reticulum retention signal
F:344/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 51.7%; Score 45; DB 2; Length 416;
Best Local Similarity 83.3%; Pred. No. 23;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPLLLGLGPAA 13
:|||||||
Db 5 VPLLGLGLGAA 16

RESULT 9
A34154
calreticulin precursor, skeletal muscle - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Sep-1999


```
C/Accession: A34154; S13047
R/Fliegel, L.; Burns, K.; MacLennan, D.H.; Reithmeier, R.A.F.; Michalak, M.
J. Biol. Chem. 264, 21522-21528, 1989
A/Title: Molecular cloning of the high affinity calcium-binding protein (calreticulin)
A/Reference number: A34154; MUID:90094320; PMID:2600080
A/Accession: A34154
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-418 <FLI>
A/Cross-references: GB:J05138; NID:g164858; PIDN:AA31188.1; PID:g164859
R/Treves, S.; de Mattei, M.; Ianfredi, M.; Villa, A.; Green, N.M.; MacLennan, D.H.; Meld
Biochem. J. 271, 473-480, 1990
A/Title: Calreticulin is a candidate for a calsequestrin-like function in Ca(2+)-storage
A/Reference number: S13045; MUID:91054414; PMID:2241926
A/Accession: S13047
A/Molecule type: protein
A/Residues: 19-32 <TR>
C/Superfamily: calreticulin
C/Keywords: skeletal muscle
F/1-17/Domain: signal sequence #status predicted <SIG>
F/415-418/Region: endoplasmic reticulum retention signal

Query Match      51.7%; Score 45; DB 1; Length 418;
Best Local Similarity 83.3%; Pred. No. 23;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 LPLLLGLGPAA 13
      :|||||:|
Db      5 VPLLLGLGLAA 16

RESULT 10
D97073
PTS system, beta-glucosides-specific IIABC component [imported] - Clostridium acetobutyli
C/Species: Clostridium acetobutylicum
C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C/Accession: D97073
R/Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
A/Reference number: A96900; MUID:21359325; PMID:21359325
A/Accession: D97073
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-628 <KUR>
A/Cross-references: GB:AE001437; PIDN:AAK79375.1; PID:g15024346; GSPDB:GN00168
A/Experimental source: Clostridium acetobutylicum ATCC824
C/Genetics:
A/Gene: CAC1407
C/Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase

Query Match      51.7%; Score 45; DB 2; Length 628;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY      1 MLPLLLGLGPACWA 16
      |:|||:|:|:|
Db      264 MVPLSLVIGPISTWA 279

RESULT 11
A55429
11-cis retinol dehydrogenase (EC 1.1.1.-) - bovine
C/Species: Bos primigenius taurus (cattle)
C/Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 29-Sep-1999
C/Accession: A55429
R/Simon, A.; Hellman, U.; Wernstedt, C.; Eriksson, U.
J. Biol. Chem. 270, 1107-1112, 1995
A/Title: The retinal pigment epithelial-specific 11-cis retinol dehydrogenase belongs to
A/Reference number: A55429; MUID:95138097; PMID:7836368
A/Accession: A55429
A/Status: preliminary; not compared with conceptual translation
```

```
A/Molecule type: mRNA
A/Residues: 1-318 <SIM>
A/Cross-references: GB:X82262; NID:g663170; PIDN:CAA57715.1; PID:g663171
C/Superfamily: retinol dehydrogenase; short-chain alcohol dehydrogenase homology
C/Keywords: membrane protein; NAD; oxidoreductase
F/29-206/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match      50.6%; Score 44; DB 2; Length 318;
Best Local Similarity 71.4%; Pred. No. 25;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2 LPLLLGLGPACW 15
      :|||||:|
Db      3 LPLLLGVLLMAALW 16

RESULT 12
F75407
probable transport protein - Deinococcus radiodurans (strain R1)
C/Species: Deinococcus radiodurans
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C/Accession: F75407
R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A/Reference number: A75250; MUID:20036896; PMID:10567266
A/Accession: F75407
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-553 <WHI>
A/Cross-references: GB:AE001980; GB:AE000513; NID:g6459086; PIDN:AAF10908.1; PID:g645908;
A/Experimental source: strain R1
C/Genetics:
A/Gene: DR1336
A/Map.position: 1

Query Match      50.6%; Score 44; DB 2; Length 553;
Best Local Similarity 58.3%; Pred. No. 42;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      4 LLLGLGPACW 15
      :|||:|:|
Db      208 LLLGVMGMTLCW 219

RESULT 13
T18291
patched protein - zebra fish
C/Species: Brachydanio rerio (zebra fish)
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C/Accession: T18291
R/Concorde, J.P.; Lewis, K.E.; Moore, J.W.; Goodrich, L.V.; Johnson, R.L.; Scott, M.P.;
Development 122, 2835-2846, 1996
A/Title: Spatial regulation of a zebrafish patched homologue reflects the roles of sonic
A/Reference number: Z18860; MUID:96379744; PMID:8787757
A/Accession: T18291
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1220 <CON>
A/Cross-references: EMBL:X98883; PIDN:CAA67386.1
C/Genetics:
A/Gene: ptc1
C/Superfamily: Drosophila membrane protein patched

Query Match      50.6%; Score 44; DB 2; Length 1220;
Best Local Similarity 61.5%; Pred. No. 86;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      1 MLPLLLGLGPAA 13
      :||:|:|:|
Db      1159 LPLVLLSLMGPPA 1171
```

RESULT 14

C81919
probable membrane protein NMA0754 [imported] - Neisseria meningitidis (strain Z2491 sero
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: C81919
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: C81919
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-119 <PAR>
A:Cross-references: GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CAB84037.1; PID:g737947
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA0754

Query Match 49.4%; Score 43; DB 2; Length 119;
Best Local Similarity 34.5%; Pred. No. 15;
Matches 10; Conservative 3; Mismatches 2; Indels 14; Gaps 1;

QY 2 LPLLLGLG-----PACWA 16
: ||||:| ||||
Db 12 ISLLGIIGIFLPLPTPFVLLSACWA 40

RESULT 15

F81182
conserved hypothetical protein NMB0571 [imported] - Neisseria meningitidis (strain MC58
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: F81182
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizzi, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: F81182
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-119 <TET>
A:Cross-references: GB:AE002413; GB:AE002098; NID:g7225797; PIDN:AAF40999.1; PID:g722579
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0571

Query Match 49.4%; Score 43; DB 2; Length 119;
Best Local Similarity 34.5%; Pred. No. 15;
Matches 10; Conservative 3; Mismatches 2; Indels 14; Gaps 1;

QY 2 LPLLLGLG-----PACWA 16
: ||||:| ||||
Db 12 ISLLGIIGIFLPLPTPFVLLSACWA 40

Search completed: January 12, 2004, 08:16:49
Job time : 3.92615 secs

THIS PAGE BLANK (USPTO)

DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ephrin type-B receptor 3 precursor (EC 2.7.1.112) (Tyrosine-protein
 DE kinase receptor MDK-5) (Developmental kinase 5) (SEK-4).
 GN EPHB3 OR ETK2 OR MDK5 OR SEK4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OK NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Embryo;
 RX MEDLINE=96074837; PubMed=7478528;
 RA Ciosek T., Lerch M.M., Ullrich A.;
 RT "Cloning, characterization, and differential expression of MDK2 and
 RT MDK5, two novel receptor tyrosine kinases of the eck/eph family.";
 RL Oncogene 11:2085-2095(1995).
 RN [2]
 RP SEQUENCE OF 719-993 FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Brain;
 RX MEDLINE=95034306; PubMed=7947319;
 RA Becker N., Seltanidou T., Murphy P., Mattei M.-G., Topilko P.,
 RA Nieto A., Wilkinson D.G., Charnay P., Gilarde P.;
 RT "Several receptor tyrosine kinase genes of the Eph family are
 RT segmentally expressed in the developing hindbrain.";
 RL Mech. Dev. 47:3-17(1994).
 RN [3]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=20171264; PubMed=10704386;
 RA Imondi R., Wideman C., Kaprielian Z.;
 RT "Complementary expression of transmembrane ephrins and their receptors
 RT in the mouse spinal cord: a possible role in constraining the
 RT orientation of longitudinally projecting axons.";
 RL Development 127:1397-1410(2000).
 CC -1- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-B FAMILY.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN CELLS OF THE RETINAL GANGLION
 CC CELT LAYER DURING RETINAL AXON GUIDANCE TO THE OPTIC DISK.
 CC -1- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
 CC -1- SIMILARITY: Contains 2 fibronectin type III domains.
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. EPHRIN
 CC RECEPTOR SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; 249086; CA88910.1; -.
 DR EMBL; X76012; CA5359.1; -.
 DR PIR; I48653; I48653.
 DR PIR; I48761; I48761.
 DR HSSP; P29323; 1B4F.
 DR MGD; MGI104770; Ephb3.
 DR GO; GO:0008046; F:axon guidance receptor activity; IDA.
 DR GO; GO:0007411; P:axon guidance; IDA.
 DR InterPro; IPR001090; Ephrin_receptor.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR003962; FNIII_subd.
 DR InterPro; IPR000719; Prot_Kinase.
 DR InterPro; IPR001660; SAM.
 DR InterPro; IPR001245; Tyr_kinase.
 DR InterPro; IPR001426; YKase_receptorV.
 DR Pfam; PF01404; EPH_1bd; 1.
 DR Pfam; PF00041; fn3; 2.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00536; SAM; 1.
 DR PRINTS; PR00014; FNTYPEIII.

DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD001495; Ephrin_receptor; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00615; Eph_1bd; 1.
 DR SMART; SM00060; FN3; 2.
 DR SMART; SM00454; SAM; 1.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
 DR PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
 DR PROSITE; PS50105; SAM_DOMAIN; 1.
 KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 KW Receptor; Transmembrane; Glycoprotein; Signal; Repeat.
 FT SIGNAL 1 29
 FT CHAIN 30 993
 FT DOMAIN 30 554
 FT TRANSMEM 555 575
 FT DOMAIN 576 993
 FT DOMAIN 191 328
 FT DOMAIN 329 443
 FT DOMAIN 444 539
 FT DOMAIN 628 891
 FT DOMAIN 920 984
 FT SITE 991 993
 FT NP_BIND 634 642
 FT BINDING 660 660
 FT ACT_SITE 753 753
 FT MOD_RES 603 603
 FT MOD_RES 609 609
 FT MOD_RES 787 787
 FT MOD_RES 937 937
 FT CARBOHYD 343 343
 FT CARBOHYD 440 440
 FT CONFLICT 719 719
 SQ SEQUENCE 993 AA; 109585 MW; 0B6GA4D391266C79 CRC64;
 R -> Q (IN REF. 2).
 Query Match 65.5%; Score 57; DB 1; Length 993;
 Best Local Similarity 68.8%; Pred. No. 0.42;
 Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 MLPULLGLGPACWA 16
 Db 14 LAPLLPLLPACWA 29
 : ||||| |||||
 RESULT 3
 PC12_MOUSE STANDARD; PRT; 1180 AA.
 AC O55134;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Protocadherin 12 precursor (Vascular cadherin-2) (Vascular endothelial
 DE cadherin-2) (VE-cadherin-2) (VE-cad-2).
 GN PCDH12.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OK NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain capillary;
 RX MEDLINE=98316322; PubMed=9651350;
 RA Telo P., Breviario F., Huber P., Panzeri C., Dejana E.;
 RT "Identification of a novel cadherin (vascular endothelial cadherin-2)
 RT located at intercellular junctions in endothelial cells.";
 RL J. Biol. Chem. 273:17565-17572(1998).
 CC -1- FUNCTION: Cellular adhesion molecule that may play an important
 CC role in cell-cell interactions at interendothelial junctions.
 CC Promotes homotypic calcium dependent aggregation and adhesion and
 CC clusters at intercellular junctions. Unable to bind to catenins,

weakly associates with the cytoskeleton.
-1- SUBCELLULAR LOCATION: Type I membrane protein.
-1- TISSUE SPECIFICITY: Expressed in endothelial cells.
-1- SIMILARITY: Contains 6 cadherin domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; Y08715; CAA69965.1; -.
DR EMBL; T31066; T31066.
DR MGD; MGI:1855700; Pcdh12.
DR GO; GO:0005911; C:intercellular junction; IDA.
DR GO; GO:0008014; F:calcium-dependent cell adhesion molecule ac. . .; IDA.
DR GO; GO:0007156; P:homophilic cell adhesion; IDA.
DR InterPro; IPR002126; Cadherin.
DR Pfam; PF00028; cadherin; 5.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 6.
DR PROSITE; PS00232; CADHERIN_1; 5.
DR PROSITE; PS50268; CADHERIN_2; 6.
KW Transmembrane; Glycoprotein; Repeat; Cell adhesion; Calcium;
KW Cytoskeleton; Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 1180 PROTOCADHERIN 12.
FT DOMAIN 1 716 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 717 737 POTENTIAL.
FT DOMAIN 738 1180 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 28 135 CADHERIN 1.
FT DOMAIN 136 244 CADHERIN 2.
FT DOMAIN 245 352 CADHERIN 3.
FT DOMAIN 355 460 CADHERIN 4.
FT DOMAIN 461 565 CADHERIN 5.
FT DOMAIN 600 711 CADHERIN 6.
FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 415 415 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 582 582 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 659 659 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 662 662 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1180 AA; 128673 MW; 97927AD41FC38087 CRC64;

Query Match 56.3%; Score 49; DB 1; Length 1180;
Best Local Similarity 69.2%; Pred. No. 7.9;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MFLPLGLGPAA 13
Db 4 LLPFLGLGPFS 16

RESULT 4
CRTC_MOUSE
ID_CRTC_MOUSE STANDARD; PRT; 416 AA.
AC P14211;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Calreticulin precursor (CRP55) (Calregulin) (HACBP) (ERP60).
GN CALR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 18-48 AND 129-161.
RC STRAIN=BALB/c; TISSUE=Liver;
RX MEDLINE=90059955; PubMed=2583110;
RA Smith M.J., Koch G.L.E.;
RT "Multiple zones in the sequence of calreticulin (CRP55, calregulin,

RT HACBP), a major calcium binding ER/SR protein.";
RL EMBL J. 8:3581-3586(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93013037; PubMed=1398135;
RA Mazzarella R.A., Gold P., Cunningham M., Green M.;
RT "Determination of the sequence of an expressible cDNA clone encoding
RT ERP60/calregulin by the use of a novel nested set method.";
RL Gene 120:217-225(1992).
RN [3]
RP SEQUENCE OF 18-38.
RC TISSUE=Fibroblast;
RX MEDLINE=95009907; PubMed=7523108;
RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
RT "Separation and sequencing of familial and novel murine proteins
RT using preparative two-dimensional gel electrophoresis.";
RL Electrophoresis 15:735-745(1994).
CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC LOW AFFINITY CALCIUM-BINDING SITES.
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X14926; CAA33053.1; -.
DR EMBL; M92988; AAA37569.1; -.
DR PIR; S06763; S06763.
DR SWISS-2DPAGE; P14211; MOUSE.
DR MGD; MGI:88252; Calr.
DR InterPro; IPR001580; Calreticulin.
DR InterPro; IPR000886; ER_target.
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PRODOM; PD001866; Calreticulin; 1.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT SIGNAL 1 17
FT CHAIN 18 416 CALRETICULIN.
FT DOMAIN 18 197 N-DOMAIN.
FT DOMAIN 198 308 P-DOMAIN.
FT DOMAIN 309 416 C-DOMAIN.
FT DOMAIN 191 255 4 X APPROXIMATE REPEATS.
FT REPEAT 191 202 1-1.
FT REPEAT 210 221 1-2.
FT REPEAT 227 238 1-3.
FT REPEAT 244 255 1-4.
FT DOMAIN 259 297 3 X APPROXIMATE REPEATS.
FT REPEAT 259 269 2-1.
FT REPEAT 273 283 2-2.
FT REPEAT 287 297 2-3.
FT DOMAIN 351 407 ASP/GLU/LYS-RICH.
FT DISULFID 137 163 BY SIMILARITY.
FT SITE 413 416 PREVENT SECRETION FROM ER.
SQ SEQUENCE 416 AA; 47994 MW; 24C03B00913408D8 CRC64;

Query Match 51.7%; Score 45; DB 1; Length 416;
Best Local Similarity 83.3%; Pred. No. 13;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPPLGLGLGPAA 13
Db 5 VPLLLGLGLGAA 16

```
RESULT 5
CRTC_RAT ID CRTC_RAT STANDARD; PRT; 416 AA.
AC P18418; P10452;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Calreticulin precursor (CRP55) (Calregulin) (HACBP) (ERP60) (CALBP)
DE (Calcium-binding protein 3) (CABP3).
GN CALR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain cortex;
RX MEDLINE=90370496; PubMed=2395661;
RA Murthy K.K., Banville D., Srikanth C.B., Carrier F., Bell A.,
RA Holmes C., Patel Y.C.;
RT "Structural homology between the rat calreticulin gene product and
RT the Onchocerca volvulus antigen Ral-1."
RL Nucleic Acids Res. 18:4933-4933(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=93202172; PubMed=8453984;
RA Nakamura M., Moriya M., Baba T., Michikawa Y., Yamanobe T., Arai K.,
RA Okinaga S., Kobayashi T.;
RT "An endoplasmic reticulum protein, calreticulin, is transported into
RT the acrosome of rat sperm."
RL Exp. Cell Res. 205:101-110(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=95181573; PubMed=7876339;
RA Soenichsen B., Fuelektrug J., van Nguyen P., Diekmann W.,
RA Robinson D.G., Mieskes G.;
RT "Retention and retrieval: both mechanisms cooperate to maintain
RT calreticulin in the endoplasmic reticulum."
RL J. Cell Sci. 107:2705-2717(1994).
RN [4]
RP SEQUENCE OF 270-358 FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Lone Y.C., Bailly A., Latruffe N.;
RL Submitted (DEC-1988) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 18-29.
RX MEDLINE=91054414; PubMed=2241926;
RA Treves S., de Mattei M., Lanfredi M., Villa A., Green N.M.,
RA MacLennan D.H., Meldolesi J., Pozzan T.;
RT "Calreticulin is a candidate for a calsequestrin-like function in
RT Ca2(+)-storage compartments (calciosomes) of liver and brain."
RL Biochem. J. 271:473-480(1990).
RN [6]
RP SEQUENCE OF 18-32.
RC STRAIN=Sprague-Dawley; TISSUE=Testis;
RX MEDLINE=92360010; PubMed=1497655;
RA Nakamura M., Michikawa Y., Baba T., Okinaga S., Arai K.;
RT "Calreticulin is present in the acrosome of spermatids of rat
RT testis."
RL Biochem. Biophys. Res. Commun. 186:668-673(1992).
RN [7]
RP SEQUENCE OF 18-32.
RC STRAIN=LEC; TISSUE=Liver;
RX MEDLINE=94072621; PubMed=8251535;
RA Yokoi T., Nagayama S., Kajiwara R., Kawaguchi Y., Horiuchi R.,
RA Kametaki T.;
RT "Identification of protein disulfide isomerase and calreticulin as
RT autoimmune antigens in LEC strain of rats."
RL Biochim. Biophys. Acta 1158:339-344(1993).
CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
```

```
CC LOW AFEINITY CALCIUM-BINDING SITES.
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC -1- CAUTION: WAS ORIGINALLY (REF.2) THOUGHT TO BE D-BETA-
CC HYDROXYBUTYRATE DEHYDROGENASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D78308; BAA11345.1; -.
DR EMBL; X53363; CAA37446.1; -.
DR EMBL; X13702; CAA31987.1; ALT_SEQ.
DR EMBL; X79327; CAA55890.1; -.
DR PIR; JH0819; JH0819.
DR PDB; 1HNN; 26-FEB-02.
DR PDB; 1K91; 12-OCT-02.
DR PDB; 1K9C; 12-OCT-02.
DR InterPro; IPR001580; Calreticulin.
DR InterPro; IPR000886; ER_target.
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR ProDom; PD001866; Calreticulin; 1.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal; 3D-structure.
FT SIGNAL 1 17
FT CHAIN 18 416 CALRETICULIN.
FT DOMAIN 18 197 N-DOMAIN.
FT DOMAIN 198 308 P-DOMAIN.
FT DOMAIN 309 416 C-DOMAIN.
FT DOMAIN 191 255 4 X APPROXIMATE REPEATS.
FT REPEAT 191 202 1-1.
FT REPEAT 210 221 1-2.
FT REPEAT 227 238 1-3.
FT REPEAT 244 255 1-4.
FT DOMAIN 259 297 3 X APPROXIMATE REPEATS.
FT REPEAT 259 269 2-1.
FT REPEAT 273 283 2-2.
FT REPEAT 287 297 2-3.
FT DOMAIN 351 407 ASP/GLU/LYS-RICH.
FT DISULFID 137 163 BY SIMILARITY.
FT SITE 413 416 PREVENT SECRETION FROM ER.
SQ SEQUENCE 416 AA; 47995 MW; 2E6713CED31A2970 CRC64;
Query Match 51.7%; Score 45; DB 1; Length 416;
Best Local Similarity 83.3%; Pred. NO. 13;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 LPLLGLGPAA 13
Db 5 VPLLGLGLAA 16
RESULT 6
CRTL_BOVIN ID CRTL_BOVIN STANDARD; PRT; 417 AA.
AC P52193; Q8S053;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Calreticulin, brain isoform 1 precursor (CRP55) (Calregulin) (HACBP).
GN CALR OR CRT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
```



```
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Hossain M.A., Takuwa K., Minakata H., Nakajima T.;
RT "Bovine brain calreticulin.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 18-417.
RC TISSUE=Brain;
RX MEDLINE=94183174; PubMed=8135753;
RA Matsuo K., Seta K., Yamakawa Y., Okuyama T., Shinoda T., Ise T.;
RT "Covalent structure of bovine brain calreticulin.";
RL Biochem. J. 298:435-442(1994).
CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC -1- LOW AFFINITY CALCIUM-BINDING SITES.
CC -1- SUBUNIT: Monomer.
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB067687; BAB86913.1; -.
DR InterPro; IPR001580; Calreticulin.
DR InterPro; IPR000886; ER_target.
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR ProDom; PD001866; Calreticulin; 1.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Glycoprotein; Signal.
FT SIGNAL 1 17
FT CHAIN 18 417 CALRETICULIN, BRAIN ISOFORM 1.
FT DOMAIN 18 197 N-DOMAIN.
FT DOMAIN 198 308 P-DOMAIN.
FT DOMAIN 309 417 C-DOMAIN.
FT DOMAIN 191 255 4 X APPROXIMATE REPEATS.
FT REPEAT 191 202 1-1.
FT REPEAT 210 221 1-2.
FT REPEAT 227 238 1-3.
FT REPEAT 244 255 1-4.
FT DOMAIN 259 297 3 X APPROXIMATE REPEATS.
FT REPEAT 259 269 2-1.
FT REPEAT 273 283 2-2.
FT REPEAT 287 297 2-3.
FT DOMAIN 351 407 ASP/GLU/LYS-RICH.
FT DISULFID 137 163
FT CARBOHYD 179 179
FT SITE 414 417
SQ SEQUENCE 417 AA; 48038 MW; 7BF812C7B5417BE9 CRC64;
Query Match 51.7%; Score 45; DB 1; Length 417;
Best Local Similarity 83.3%; Pred. No. 13;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 2 LPLLGLLGPAA 13
Db 5 VPLLGLLGLAA 16
```

```
RESULT 7
CRTC_CRIGR STANDARD; PRT; 417 AA.
ID CRTC_CRIGR
AC Q8K3H7;
```

```
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, last sequence update)
DT 28-FEB-2003 (Rel. 41, last annotation update)
DE Calreticulin precursor (CRP55) (Calregulin) (HACBP) (ERP60).
GN CALR.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RA Chung J.Y., Lee G.M.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC -1- LOW AFFINITY CALCIUM-BINDING SITES.
CC -1- SUBUNIT: Monomer.
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AY100688; AAM48568.1; -.
DR InterPro; IPR001580; Calreticulin.
DR InterPro; IPR000886; ER_target.
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR ProDom; PD001866; Calreticulin; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
DR PROSITE; PS00014; ER_TARGET; 1.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Glycoprotein; Signal.
FT SIGNAL 1 17
FT CHAIN 18 417 CALRETICULIN.
FT DOMAIN 18 197 N-DOMAIN.
FT DOMAIN 198 308 P-DOMAIN.
FT DOMAIN 309 417 C-DOMAIN.
FT DOMAIN 191 255 4 X APPROXIMATE REPEATS.
FT REPEAT 191 202 1-1.
FT REPEAT 210 221 1-2.
FT REPEAT 227 238 1-3.
FT REPEAT 244 255 1-4.
FT DOMAIN 259 297 3 X APPROXIMATE REPEATS.
FT REPEAT 259 269 2-1.
FT REPEAT 273 283 2-2.
FT REPEAT 287 297 2-3.
FT DOMAIN 351 407 ASP/GLU/LYS-RICH.
FT DISULFID 137 163
FT SITE 414 417
SQ SEQUENCE 417 AA; 48242 MW; D617DA37D14F2D45 CRC64;
Query Match 51.7%; Score 45; DB 1; Length 417;
Best Local Similarity 83.3%; Pred. No. 13;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 2 LPLLGLLGPAA 13
Db 5 VPLLGLLGLAA 16
```

```
RESULT 8
CRTC_RABIT STANDARD; PRT; 418 AA.
ID CRTC_RABIT
AC P15253;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, last sequence update)
```


DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Calreticulin precursor (CRP55) (Calregulin) (HACBP) (Erp60).
 GN CALR.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Slow-twitch skeletal muscle;
 RX MEDLINE=90094320; PubMed=2600080;
 RA Fliegel L., Burns K., MacLennan D.H., Reithmeier R.A.F., Michalak M.;
 RT "Molecular cloning of the high affinity calcium-binding protein
 (calreticulin) of skeletal muscle sarcoplasmic reticulum.";
 RL J. Biol. Chem. 264:21522-21528(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fast-twitch skeletal muscle;
 RX MEDLINE=91282795; PubMed=2059224;
 RA Fliegel L., Michalak M.;
 RT "Fast-twitch and slow-twitch skeletal muscles express the same
 isoform of calreticulin.";
 RL Biochem. Biophys. Res. Commun. 177:979-984(1991).
 RN [3]
 RP SEQUENCE OF 18-36.
 RX MEDLINE=91054414; PubMed=2241926;
 RA Treves S., de Mattei M., Lanfredi M., Villa A., Green N.M.,
 RA MacLennan D.H., Meldolesi J., Pozzan T.;
 RT "Calreticulin is a candidate for a calsequestrin-like function in
 Ca2(+)-storage compartments (calciosomes) of liver and brain.";
 RL Biochem. J. 271:473-480(1990).
 RN [4]
 RP SEQUENCE OF 18-46.
 RX MEDLINE=91201375; PubMed=2016321;
 RA Milner R.E., Baksh S., Shemanko C., Carpenter M.R., Smillie L.,
 RA Vance J.E., Opas M., Michalak M.;
 RT "Calreticulin, and not calsequestrin, is the major calcium binding
 protein of smooth muscle sarcoplasmic reticulum and liver endoplasmic
 reticulum.";
 RL J. Biol. Chem. 266:7155-7165(1991).
 RN [5]
 RP PARTIAL SEQUENCE.
 RC TISSUE=Lung;
 RX MEDLINE=92002038; PubMed=1911780;
 RA Guan S., Fallick A.M., Williams D.E., Cashman J.R.;
 RT "Evidence for complex formation between rabbit lung flavin-containing
 monooxygenase and calreticulin.";
 RL Biochemistry 30:9892-9900(1991).
 CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 CC LOW AFFINITY CALCIUM-BINDING SITES.
 CC -1- SUBUNIT: Monomer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
 CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; J05138; AAA31188.1; -.
 DR PIR; A34154; A34154.
 DR InterPro; IPR001580; Calreticulin.
 DR InterPro; IPR000886; ER_target.
 DR Pfam; PF00262; calreticulin; 1.
 DR PRINTS; PR00626; CALRETICULIN.
 DR ProDom; PD001866; Calreticulin; 1.
 DR PROSITE; PS00014; ER_TARGET; 1.
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.

KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
 FT SIGNAL 1 17
 FT CHAIN 18 418
 FT DOMAIN 18 197
 FT DOMAIN 198 308
 FT DOMAIN 309 418
 FT DOMAIN 191 255
 FT REPEAT 191 202
 FT REPEAT 210 221
 FT REPEAT 227 238
 FT REPEAT 244 255
 FT DOMAIN 259 297
 FT REPEAT 259 269
 FT REPEAT 273 283
 FT REPEAT 287 297
 FT DOMAIN 351 408
 FT DISULFID 137 163
 FT SITE 415 418
 FT VARIANT 35 35
 FT CONFLICT 90 90
 SQ SEQUENCE 418 AA; 48275 MW; B6082B689DC763A6 CRC64;
 Query Match 51.7%; Score 45; DB 1; Length 418;
 Best Local Similarity 83.3%; Pred. No. 13;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 LPLLGLGPAA 13
 Db 5 VPLLGLGLAA 16
 RESULT 9
 ENV_SRV2
 ID ENV_SRV2 STANDARD; PRT; 574 AA.
 AC P51515;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE ENV polypeptide (Coat polypeptide) [Contains: Coat protein GP70; Coat
 DE protein GP20].
 GN ENV.
 OS Simian retrovirus SRV-2.
 OC Viruses; Retroid viruses; Retroviridae; Betaretrovirus.
 OX NCBI_TaxID=39068;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87151131; PubMed=2435057;
 RA Thayer R.M., Power M.D., Bryant M.L., Gardner M.B., Barr P.J.,
 RA Luciw P.A.;
 RT "Sequence relationships of type D retroviruses which cause simian
 RT acquired immunodeficiency syndrome.";
 RL Virology 157:317-329(1987).
 CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M16605; AAA47563.1; ALT_INIT.
 DR HSSP; P03385; IMOF.
 DR InterPro; IPR002050; Env_polypeptide.
 DR Pfam; PF00429; ENV_polypeptide; 1.
 KW Coat protein; Glycoprotein; Polypeptide; Transmembrane.
 FT CHAIN 1 382
 FT CHAIN 383 574
 FT TRANSMEM 386 406
 FT TRANSMEM 515 535
 FT CARBOHYD 117 117
 FT CARBOHYD 233 233
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).

```
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 312 312 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 318 318 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 475 475 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 574 AA; 62952 MM; 662B1E11A437E055 CRC64;

Query Match 51.7%; Score 45; DB 1; Length 574;
Best Local Similarity 64.3%; Pred. No. 17;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy 1 MLPLLGLGPPAC 14
Db 510 LLPYLLPLGLPFC 523

RESULT 10
RDH1_BOVIN STANDARD; PRT; 318 AA.
AC Q27979;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 11-cis retinol dehydrogenase (EC 1.1.1.105) (11-cis RDH) (P32).
GN RDH5 OR RDH1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RT TISSUE=Eye;
RX MEDLINE=95138097; PubMed=7836368;
RA Simon A., Hellman U., Wernstedt C., Eriksson U.;
RT "The retinal pigment epithelial-specific 11-cis retinol dehydrogenase
RT belongs to the family of short chain alcohol dehydrogenases.";
RL J. Biol. Chem. 270:1107-1112(1995).
CC -1- FUNCTION: STEREOSPECIFIC 11-CIS RETINOL DEHYDROGENASE, WHICH
CC CATALYZES THE FINAL STEP IN THE BIOSYNTHESIS OF 11-CIS
CC RETINALDEHYDE, THE UNIVERSAL CHROMOPHORE OF VISUAL PIGMENTS.
CC ACTIVE IN THE PRESENCE OF NAD AS COFACTOR BUT NOT IN THE PRESENCE
CC OF NADP.
CC -1- CATALYTIC ACTIVITY: Retinol + NAD(+) = retinal + NADH.
CC -1- PATHWAY: Retinoic acid biosynthesis; first (rate-limiting) step.
CC -1- SUBCELLULAR LOCATION: Membrane-associated.
CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC (SDR) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X82262; CAA57715.1; -.
DR PIR; A55429; A55429.
DR HSSP; P14061; 1FDU.
DR InterPro; IPR002198; ADH_short.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; FALSE_NEG.
KW Oxidoreductase; NAD; Membrane; Vision.
FT NP_BIND 32 56 NADP (BY SIMILARITY).
FT ACT_SITE 175 175 BY SIMILARITY.
SQ SEQUENCE 318 AA; 35036 MW; 78B4A1E43B5FB351 CRC64;
```

```
Query Match 50.6%; Score 44; DB 1; Length 318;
Best Local Similarity 71.4%; Pred. No. 15;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 2 LPPLLGLGPPACW 15
Db 3 LPPLLGLVLMALW 16

RESULT 11
PC12_HUMAN STANDARD; PRT; 1184 AA.
ID PC12_HUMAN
AC Q9NPG4; Q96KB8; Q9H7Y6; Q9H8E0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE protocadherin 12 precursor (Vascular cadherin-2) (Vascular endothelial
DE cadherin-2) (VE-cadherin-2) (VE-cad-2).
GN PCDH12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20202599; PubMed=10716726;
RA Wu Q., Maniatis T.;
RT "Large exons encoding multiple ectodomains are a characteristic
RT feature of protocadherin genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:3124-3129(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal kidney;
RX MEDLINE=20515266; PubMed=11063261;
RA Ludwig D., Lorenz J., Dejane E., Bohlén P., Hicklin D.J., Witte L.,
RA Pytowski B.;
RT "cDNA cloning, chromosomal mapping, and expression analysis of human
RT VE-cadherin-2.";
RL Mamm. Genome 11:1030-1033(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Seki N., Hattori A., Hayashi A., Kozuma S., Muramatsu M., Saito T.;
RT "Human vascular cadherin-2.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo, and placenta;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "MEDO human cDNA sequencing project.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Cellular adhesion molecule that may play an important
CC role in cell-cell interactions at interendothelial junctions.
CC Promotes homotypic calcium dependent aggregation and adhesion and
CC clusters at intercellular junctions. Unable to bind to catenins,
CC weakly associates with the cytoskeleton (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in highly vascularized tissues
CC including the heart and placenta, but most tissues contain a low
CC level of expression. Prominent expression in the spleen.
CC -1- SIMILARITY: Contains 6 cadherin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF231025; AAF61931.1; -.
DR EMBL; AF240635; AAF73962.1; -.
DR EMBL; AB026893; BAA95162.1; -.
DR EMBL; AK024140; BAB14837.1; -.
DR EMBL; AK023785; BAB14677.1; -.
DR EMBL; AK027282; BAB55016.1; -.
DR Genew; HGNC:8657; PCDH12.
DR MIM; 605622; -.
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0008038; P:neuronal cell recognition; TAS.
DR InterPro; IPR002126; Cadherin.
DR Pfam; PF00028; cadherin; 5.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 5.
DR PROSITE; PS00232; CADHERIN_1; 5.
DR PROSITE; PS50268; CADHERIN_2; 6.
KW Transmembrane; Glycoprotein; Repeat; Cell adhesion; Calcium;
KW Cytoskeleton; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 1184 PROTOCADHERIN 12.
FT DOMAIN 1 718 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 719 739 POTENTIAL.
FT DOMAIN 740 1184 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 28 135 CADHERIN 1.
FT DOMAIN 136 244 CADHERIN 2.
FT DOMAIN 245 352 CADHERIN 3.
FT DOMAIN 355 460 CADHERIN 4.
FT DOMAIN 461 565 CADHERIN 5.
FT DOMAIN 600 711 CADHERIN 6.
FT DOMAIN 1177 1181 POLY-SER.
FT CARBOHYD 415 415 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 582 582 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 659 659 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 662 662 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 385 385 H -> N (IN REF. 4; BAB14837).
FT CONFLICT 386 386 HN -> KD (IN REF. 4; BAB55016).
FT CONFLICT 389 390 VH -> LG (IN REF. 4; BAB55016).
FT CONFLICT 442 442 A -> V (IN REF. 4; BAB14677).
FT CONFLICT 640 640 S -> N (IN REF. 4; BAB14677).
FT CONFLICT 753 753 R -> W (IN REF. 4; BAB14677).
FT CONFLICT 814 814 A -> T (IN REF. 4; BAB14837).
FT CONFLICT 970 970 H -> Y (IN REF. 4; BAB14837).
FT CONFLICT 1051 1051 S -> C (IN REF. 4; BAB14677).
FT CONFLICT 1181 1181 S -> SSSS (IN REF. 4; BAB14837).
SQ SEQUENCE 1184 AA; 128993 MW; 45314473DC503E8D CRC64;

Query Match 50.6%; Score 44; DB 1; Length 1184;
Best Local Similarity 81.8%; Pred. No. 45;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MUPLLGLGP 11
Db 4 LQLLLGLGP 14

RESULT 12
PTC1_BRARE STANDARD; PRT; 1220 AA.
ID PTC1_BRARE
AC Q98864;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Patched protein homolog 1 (Patched 1) (PTC1).
GN PTC1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC TISSUE=Embryo;
RX MEDLINE=96379744; PubMed=8787757;
RA Concorde J.-P., Lewis K.E., Moore J.W., Goodrich L.V., Johnson R.L.,
RA Scott M.P., Ingham P.W.;
RT "Spatial regulation of the zebrafish patched homologue reflects the
RT roles of sonic hedgehog and protein kinase A in neural tube and somite
RT patterning.";
RL Development 122:2835-2846(1996).
CC -!- FUNCTION: ACTS AS A RECEPTOR FOR SONIC HEDGEHOG (SHH), INDIAN
CC HEDGEHOG (IHH) AND DESERT HEDGEHOG (DHH). ASSOCIATES WITH THE
CC SMOOTHENED PROTEIN (SMO) TO TRANSDUCE THE HEDGEHOG'S PROTEINS
CC SIGNAL (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: DETECTED IN EMBRYONIC PRESOMITIC MESODERM,
CC NEUROECTODERM, TISSUE SURROUNDING THE NOTOCHORD, VENTRAL NEURAL
CC TUBE.
CC -!- DEVELOPMENTAL STAGE: AT ALL STAGES, EXPRESSION CORRESPONDS TO THE
CC LOCALIZATION OF SHH. FIRST DETECTED DURING GASTRULATION. BY 36
CC HOURS, PTC1 APPEARS IN THE FIRST BRANCHIAL ARCH AND THE POSTERIOR
CC MESENCHYME OF THE FIN BUD; BY 48 HOURS, IN THE HINDBRAIN AND
CC FOREGUT.
CC -!- INDUCTION: ACTIVATED BY SONIC HEDGEHOG.
CC -!- PTM: GLYCOSYLATION IS NECESSARY FOR SHH BINDING (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE PATCHED FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X98883; CAA67386.1; -.
DR PIR; T18291; T18291.
DR ZFIN; ZDB-GENE-980526-44; ptc1.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR004766; Patched_rec.
DR Pfam; PF02460; Patched_1.
DR TIGRfam; TIGR00918; 2A060602; 1.
DR PROSITE; PS50156; SSD; 1.
KW Receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 84 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 85 105 POTENTIAL.
FT DOMAIN 106 419 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 420 440 POTENTIAL.
FT DOMAIN 441 449 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 450 470 POTENTIAL.
FT DOMAIN 471 484 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 485 505 POTENTIAL.
FT DOMAIN 506 528 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 529 549 POTENTIAL.
FT DOMAIN 550 558 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 559 579 POTENTIAL.
FT DOMAIN 580 739 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 740 760 POTENTIAL.
FT DOMAIN 761 1016 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1017 1037 POTENTIAL.
FT DOMAIN 1038 1044 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 1045 1065 POTENTIAL.
FT DOMAIN 1066 1072 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1073 1093 POTENTIAL.
FT DOMAIN 1094 1110 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 1111 1131 POTENTIAL.
FT DOMAIN 1132 1143 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1144 1164 POTENTIAL.
FT DOMAIN 1165 1220 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 865 865 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 888 888 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1220 AA; 135544 MW; D10A9D04115F532D CRC64;
```


Query Match 50.6%; Score 44; DB 1; Length 1220;
Best Local Similarity 61.5%; Pred. No. 46;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 MLPLLLGLGPAA 13
:|:|:|:|:|:|
Db 1159 LLPVLLSLMGPPA 1171

RESULT 13
TAM_PSEAE STANDARD; PRT; 275 AA.
ID TAM_PSEAE
AC Q9I0S1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Trans-aconitate 2-methyltransferase (EC 2.1.1.144).
GN TAM OR PA2564.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
CC -|- FUNCTION: Catalyzes the S-adenosylmethionine monomethyl
CC esterification of trans-aconitate (By similarity).
CC -|- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + trans-aconitate = S-
CC adenosyl-L-homocysteine + (E)-3-(methoxycarbonyl)pent-2-enedioate.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- SIMILARITY: BELONGS TO THE METHYLTRANSFERASE SUPERFAMILY. TAM
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC EMBL; AE004684; AAG05952.1; -.
DR PIR; H83325; H83325.
DR HAMAP; MF_00560; -; 1.
DR InterPro; IPR000051; SAM bind.
KW Transferase; Methyltransferase; Complete proteome.
SQ SEQUENCE 275 AA; 31674 MM; A902E2B71FF2E06B CRC64;

Query Match 49.4%; Score 43; DB 1; Length 275;
Best Local Similarity 56.2%; Pred. No. 18;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 1 MLPLLLGLGPACWA 16
:|:|:|:|:|:|
Db 124 LLPRLTGLTPGCTA 139

RESULT 14
MIAA_BUCBP STANDARD; PRT; 323 AA.
ID MIAA_BUCBP
AC P59507;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE tRNA delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.8) (IPP
DE transferase) (Isopentenyl-diphosphate:tRNA isopentenyltransferase)
DE (IPTase) (IPPT).
GN MIAA OR BHP514.
OS Buchnera aphidicola (subsp. Baizongia pistaciae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=135842;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426901; PubMed=12522265;
RA Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,
RA Bastolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J.,
RA Tamames J., Viguera E., Latorre A., Valencia A., Moran F., Moya A.;
RT "Reductive genome evolution in Buchnera aphidicola."
RL Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).
CC -|- FUNCTION: Catalyzes the first step in the biosynthesis of 2-
CC methylthio-N6-(delta(2)-isopentenyl)-adenosine (MS[2]I[6]A)
CC adjacent to the anticodon of several tRNA species (By similarity).
CC -|- CATALYTIC ACTIVITY: Isopentenyl diphosphate + tRNA = diphosphate +
CC tRNA containing 6-isopentenyladenosine.
CC -|- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC EMBL; AE014017; AAO27217.1; -.
DR HAMAP; MF_00185; -; 1.
KW Transferase; Nucleotidyltransferase; tRNA processing; ATP-binding;
KW Complete proteome.
FT NP BIND 21 ATP (POTENTIAL).
SQ SEQUENCE 323 AA; 37553 MM; D60CFE62D90E1C9D CRC64;

Query Match 49.4%; Score 43; DB 1; Length 323;
Best Local Similarity 58.3%; Pred. No. 21;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 3 PLILGLGPAA 14
||:|:|:|:|
Db 14 PLVFLMGPTAC 25

RESULT 15
COA2_POVMK STANDARD; PRT; 341 AA.
ID COA2_POVMK
AC P24596;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Coat protein VP2 (Contains: Coat protein VP3).
OS Mouse polyomavirus (strain Kilham).
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10638;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91196237; PubMed=1849675;
RA Mayer M., Doerries K.;
RT "Nucleotide sequence and genome organization of the murine
RT polyomavirus, Kilham strain."
RL Virology 181:469-480(1991).
RN [2]
RP REVISIONS TO 58 AND 121.
RA Mayer M., Doerries K.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
CC -|- ALTERNATIVE PRODUCTS:
CC Event=Alternative initiation;
CC Comment=2 isoforms, VP2 (shown here) and VP3, are produced by

CC alternative initiation;
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M55904; AAA46554.1; -.
 DR EMBL; M55904; AAA46555.1; -.
 DR InterPro; IPR001070; Polyoma_coat2.
 DR Pfam; PF00761; Polyoma_coat2; 1.
 KW Late protein; Coat protein; Alternative initiation.
 FT CHAIN 1 341 COAT PROTEIN VP2, ISOFORM VP2.
 FT CHAIN 120 341 COAT PROTEIN VP2, ISOFORM VP3.
 FT INIT MET 120 120 FOR ISOFORM VP3.
 FT CONFLICT 58 58 R -> A (IN REF. 1).
 FT CONFLICT 121 121 A -> V (IN REF. 1).
 SQ SEQUENCE 341 AA; 37396 MW; 9F8E3D25514BBF75 CRC64;

Query Match 49.4%; Score 43; DB 1; Length 341;
 Best Local Similarity 90.0%; Pred. No. 22;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MLPLLGLLG 10
 |||||
 Db 283 MLPLLGLYG 292

Search completed: January 12, 2004, 08:13:01
 Job time : 2.57485 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:10 ; Search time 2.20359 Seconds
(without alignments)
1873.686 Million cell updates/sec

Title: US-10-085-167-2_COPY_1_16
Perfect score: 87
Sequence: 1 MLPLLLGLGPACWA 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------|--------------------|
| 1 | 87 | 100.0 | 329 4 | Q8IV25 | Q8iv25 homo sapien |
| 2 | 73 | 83.9 | 326 11 | Q8R066 | Q8r066 mus musculu |
| 3 | 57 | 65.5 | 993 11 | Q91YS9 | Q91ys9 mus musculu |
| 4 | 52 | 59.8 | 115 16 | Q51537 | Q51537 pseudomonas |
| 5 | 49 | 56.3 | 155 16 | Q8PCR0 | Q8pcr0 xanthomonas |
| 6 | 49 | 56.3 | 998 11 | Q8C883 | Q8c883 mus musculu |
| 7 | 49 | 56.3 | 1061 5 | Q9V9L4 | Q9v9l4 drosophila |
| 8 | 48 | 55.2 | 379 11 | Q9DCW1 | Q9dcw1 mus musculu |
| 9 | 47 | 54.0 | 569 8 | Q94RE8 | Q94re8 lithobius f |
| 10 | 46.5 | 53.4 | 719 16 | Q8P3B9 | Q8p3b9 xanthomonas |
| 11 | 46 | 52.9 | 228 16 | Q9AK63 | Q9ak63 streptomyce |
| 12 | 46 | 52.9 | 496 5 | Q967F2 | Q967f2 caenorhabdi |
| 13 | 46 | 52.9 | 517 11 | Q8CHQ8 | Q8chg8 mus musculu |
| 14 | 46 | 52.9 | 517 11 | Q8BJ48 | Q8bj48 mus musculu |
| 15 | 46 | 52.9 | 558 3 | Q9UVM5 | Q9uvm5 coprinus ci |
| 16 | 46 | 52.9 | 580 15 | Q9WPZ9 | Q9wpz9 simian retr |

| | | | | | | |
|----|------|------|------|----|--------|---------------------|
| 17 | 45 | 51.7 | 309 | 16 | Q8UY5 | Q8uy5 agrobacteri |
| 18 | 45 | 51.7 | 315 | 16 | Q9RU53 | Q9ru53 deinococcus |
| 19 | 45 | 51.7 | 347 | 4 | Q9Y6U7 | Q9y6u7 homo sapien |
| 20 | 45 | 51.7 | 395 | 16 | Q8PF13 | Q8pf13 xanthomonas |
| 21 | 45 | 51.7 | 488 | 5 | Q9NI45 | Q9ni45 trypanosoma |
| 22 | 45 | 51.7 | 602 | 5 | Q9VUJ3 | Q9vuj3 drosophila |
| 23 | 45 | 51.7 | 628 | 16 | Q97J79 | Q97j79 clostridium |
| 24 | 44.5 | 51.1 | 232 | 12 | Q89737 | Q89737 polyomaviru |
| 25 | 44.5 | 51.1 | 253 | 16 | Q8R9M1 | Q8r9m1 thermocanaer |
| 26 | 44 | 50.6 | 101 | 5 | Q8IF11 | Q8if11 trypanosoma |
| 27 | 44 | 50.6 | 175 | 10 | Q8GSU7 | Q8gsu7 coffea arab |
| 28 | 44 | 50.6 | 177 | 16 | Q9RKE7 | Q9rke7 streptomyce |
| 29 | 44 | 50.6 | 297 | 10 | Q8H527 | Q8h527 oryza sativ |
| 30 | 44 | 50.6 | 517 | 2 | Q9ZGB6 | Q9zgb6 streptomyce |
| 31 | 44 | 50.6 | 521 | 16 | Q8ETN6 | Q8etn6 oceanobacil |
| 32 | 44 | 50.6 | 553 | 16 | Q9RUP7 | Q9rup7 deinococcus |
| 33 | 44 | 50.6 | 1187 | 4 | Q8IUP8 | Q8iup8 homo sapien |
| 34 | 44 | 50.6 | 1413 | 13 | Q9DEF3 | Q9def3 xenopus lae |
| 35 | 43 | 49.4 | 74 | 12 | Q84349 | Q84349 simian viru |
| 36 | 43 | 49.4 | 100 | 4 | Q9UGS8 | Q9ugs8 homo sapien |
| 37 | 43 | 49.4 | 119 | 16 | Q9KOM1 | Q9kom1 neisseria m |
| 38 | 43 | 49.4 | 119 | 16 | Q9JVP6 | Q9jvp6 neisseria m |
| 39 | 43 | 49.4 | 159 | 16 | Q9RY14 | Q9ry14 deinococcus |
| 40 | 43 | 49.4 | 213 | 16 | Q8NMT4 | Q8nmt4 corynebacte |
| 41 | 43 | 49.4 | 225 | 12 | Q91NM4 | Q91nm4 polyomaviru |
| 42 | 43 | 49.4 | 225 | 12 | Q91NN7 | Q91nn7 polyomaviru |
| 43 | 43 | 49.4 | 225 | 12 | Q8JUB7 | Q8jub7 polyomaviru |
| 44 | 43 | 49.4 | 225 | 12 | Q91OS1 | Q91os1 polyomaviru |
| 45 | 43 | 49.4 | 225 | 12 | P90328 | P90328 polyomaviru |

ALIGNMENTS

RESULT 1
Q8IV25 PRELIMINARY; PRT; 329 AA.
ID Q8IV25
AC Q8IV25;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Similar to C1q and tumor necrosis factor related protein 4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC035628; AAH35628.1; -.
SQ SEQUENCE 329 AA; 35256 MW; 16064DA8182A6732 CRC64;

Query Match 100.0%; Score 87; DB 4; Length 329;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPLLLGLGPACWA 16
Db 1 MLPLLLGLGPACWA 16

RESULT 2
Q8R066 PRELIMINARY; PRT; 326 AA.
ID Q8R066
AC Q8R066;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Similar to C1q and tumor necrosis factor related protein 4.
GN 0710001E10RIK.
OS Mus musculus (Mouse).

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027315; AAH27315.1; -.
DR MGD; MGI:1914695; 0710001E10R1K.
DR InterPro; IPR001073; Clq.
DR Pfam; PF00386; Clq; 2.
DR SMART; SM00110; Clq; 2.
DR PROSITE; PS01113; Clq; 2.
SQ SEQUENCE 326 AA; 35057 MW; 72339172B7B1051A CRC64;

Query Match      83.9%; Score 73; DB 11; Length 326;
Best Local Similarity 87.5%; Pred. No. 0.0032;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLPLLGLGPAACWA 16
   |||||
Db 1 MLPLLGLGPAACWA 16

RESULT 3
Q91YS9 PRELIMINARY; PRT; 993 AA.
ID Q91YS9
AC Q91YS9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to Ephb3 (EC 2.7.1.112) (Ephrin receptor) (Tyrosine-protein
DE kinase receptor).
GN EPHB3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC TYROSINE PHOSPHATE.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. EPHRIN
CC RECEPTOR SUBFAMILY.
CC EMBL; BC014822; AAH14822.1; -.
DR MGD; MGI:104770; Ephb3.
DR InterPro; IPR001090; Ephrin_receptor.
DR InterPro; IPR003962; FNIII_subd.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR001426; YKase_receptorV.
DR Pfam; PF01404; EPH_1bd; 1.
DR Pfam; PF00041; fn3; 2.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00536; SAM; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD0001495; Ephrin_receptor; 1.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00615; EPH_1bd; 1.
DR SMART; SM00060; FN3; 2.
DR SMART; SM00454; SAM; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
```

```
DR PROSITE; PS50105; SAM_DOMAIN; 1.
KW ATP-binding; Glycoprotein; Kinase; Phosphorylation; Receptor; Repeat;
KW Transferrase; Transmembrane; Tyrosine-protein kinase.
SQ SEQUENCE 993 AA; 109661 MW; 07C6EF2AC98AE1B4 CRC64;

Query Match      65.5%; Score 57; DB 11; Length 993;
Best Local Similarity 68.8%; Pred. No. 2.4;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MLPLLGLGPAACWA 16
   : |||||
Db 14 LAPLLPLLPAGCWA 29

RESULT 4
Q51537 PRELIMINARY; PRT; 115 AA.
ID Q51537
AC Q51537;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Type 4 fimbrial biogenesis protein PILY2.
DE PILY2 OR PA4555.
GN Pseudomonas aeruginosa.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=95157261; PubMed=7854130;
RA Alm R.A.; Mattick J.S.;
RT "The pili gene product of Pseudomonas aeruginosa, required for pilus
RT biogenesis, shares amino acid sequence identity with the N-termini of
RT type 4 prepilin proteins."
RL Mol. Microbiol. 13:973-985(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=96037793; PubMed=7565109;
RA Alm R.A.; Mattick J.S.;
RT "Identification of a gene, pilV, required for type 4 fimbrial
RT biogenesis in Pseudomonas aeruginosa, whose product possesses a pre-
RT pilin-like leader sequence."
RL Mol. Microbiol. 16:485-496(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RA Alm R.A.; Hallinan J.P.; Watson A.A.; Mattick J.S.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K.; Pham X.-Q.T.; Erwin A.L.; Mizoguchi S.D.; Warrenner P.;
RA Hickey M.J.; Brinkman F.S.L.; Hufnagle W.O.; Kowalik D.J.; Lagrou M.;
RA Garber R.L.; Goltzy L.; Tolentino E.; Westbrook-Wadman S.; Yuan Y.;
RA Brody L.L.; Coulter S.N.; Folger K.R.; Kas A.; Larbig K.; Lim R.M.;
RA Smith K.A.; Spencer D.H.; Wong G.K.-S.; Wu Z.; Paulsen I.T.;
RA Reizer J.; Saier M.H.; Hancock R.E.W.; Lory S.; Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RT Nature 406:959-964(2000).
RL Nature 406:959-964(2000).
DR EMBL; L76605; AAA93503.1; -.
DR EMBL; AE004869; AAG07943.1; -.
KW Complete proteome.
SQ SEQUENCE 115 AA; 12986 MW; AC19C31640473181 CRC64;

Query Match      59.8%; Score 52; DB 16; Length 115;
Best Local Similarity 56.2%; Pred. No. 1.9;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MLPLLGLGPAACWA 16
```

Db :||:|||||
3 VLPMLLALAVPGLCWA 18

RESULT 5

Q8PCRO PRELIMINARY; PRT; 155 AA.
AC Q8PCRO;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Hypothetical protein XCC0653.
GN XCC0653.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Canavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
DR EMBL; AE012163; AAM39969.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 155 AA; 17448 MW; DEA92F2FA2C63A51 CRC64;

Query Match 56.3%; Score 49; DB 16; Length 155;
Best Local Similarity 60.0%; Pred. No. 7.1;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 MLPLLGLLGPACW 15
Db 91 MRPWRLSLTGPQACW 105

RESULT 6

Q8C883 PRELIMINARY; PRT; 998 AA.
AC Q8C883;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Protocadherin 12.
GN Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).

DR EMBL; AK048100; BAC33242.1; -
SQ SEQUENCE 998 AA; 109373 MW; 604080E3D5C37F78 CRC64;

Query Match 56.3%; Score 49; DB 11; Length 998;
Best Local Similarity 69.2%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLPLLGLLGPAA 13
Db 4 LPPFLGLLGP GS 16

RESULT 7

Q9V9L4 PRELIMINARY; PRT; 1061 AA.
AC Q9V9L4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE CG11212 protein.
GN CG11212.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktarglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclele J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003784; AAF57274.1; -
DR FlyBase; FBgn0033068; CG11212.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR006162; Ppatcne_attach.
DR InterPro; IPR000731; SSD_STM.

DR Pfam; PF02460; Patched; 1.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
DR PROSITE; PS50156; SSD; 1.
SQ SEQUENCE 1061 AA; 120880 MW; DC6F32237EBE9A96 CRC64;

Query Match 56.3%; Score 49; DB 5; Length 1061;
Best Local Similarity 57.1%; Pred. No. 41;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLPLLLGLGPAC 14
:|||||:|
Db 740 LLPVLISLFGPGSC 753

RESULT 8

Q9DCW1 PRELIMINARY; PRT; 379 AA.
AC Q9DCW1;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE 0610009J22Rik protein.
GN 0610009J22Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Wittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AK002414; BAB22082.1; -.
DR HSSP; P28990; 1CHC.
DR MGD; MGI:1918923; 0610009J22Rik.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
KW Metal-binding; Zinc; Zinc_finger.
SQ SEQUENCE 379 AA; 41577 MW; DA36AA22623E3901 CRC64;

Query Match 55.2%; Score 48; DB 11; Length 379;
Best Local Similarity 69.2%; Pred. No. 23;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLPLLLGLGPAA 13
:|||||:|
Db 31 LLPLWLGLMGPAA 43

RESULT 9

Q94RE8

ID Q94RE8 PRELIMINARY; PRT; 569 AA.
AC Q94RE8;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE NADH dehydrogenase subunit 5 (Fragment).
GN ND5.
OS Lithobius forficatus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Myriapoda; Chilopoda;
OC Pleurostigmophora; Lithobiomorpha; Lithobiidae; Lithobius.
OX NCBI_TaxID=7552;
RN [1]

RP SEQUENCE FROM N.A.
RA Hwang U.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21441907; PubMed=11557978;
RA Hwang U., Friedrich M., Choe C., Kim W.;
RT "Mitochondrial protein phylogeny joins myriapods with chelicerates."
RL Nature 413:154-157(2001).
DR EMBL; AJ270997; CAC69943.1; -.
DR InterPro; IPR003916; NADhub_oxred5.
DR InterPro; IPR001750; Oxidored_q1.
DR InterPro; IPR001516; Oxidored_q1_N.
DR InterPro; IPR006162; Pantane_attach.
DR Pfam; PF00361; oxidored_q1; 1.
DR Pfam; PF00662; oxidored_q1_N; 1.
DR PRINTS; PR01434; NADHDGNASES.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
KW NAD; Oxidoreductase; Ubiquinone; Mitochondrion.
FT NON TER 569 569
SQ SEQUENCE 569 AA; 62591 MW; 4C3104BFEA6F72AB CRC64;

Query Match 54.0%; Score 47; DB 8; Length 569;
Best Local Similarity 58.3%; Pred. No. 47;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 LLLGLGPACW 15
|:|:|:|
Db 169 LSFGLVGPVSCW 180

RESULT 10
Q8P3B9 PRELIMINARY; PRT; 719 AA.
AC Q8P3B9;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Transglycolase.
GN XCC4157.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Clapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,

RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kltajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AE012541; AAM43373.1; -
DR InterPro; IPR001823; Ald1_epimerase.
DR Pfam; PF01263; Aldose_epim; 1.
KW Complete proteome.
SQ SEQUENCE 719 AA; 77224 MW; 45F5583D9E83D132 CRC64;

Query Match 53.4%; Score 46.5; DB 16; Length 719;
Best Local Similarity 73.3%; Pred. No. 70;
Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 2 LPPLLGLGPACWA 16
Db 8 LPPLLGLGPACWA 21

RESULT 11
Q9AK63 PRELIMINARY; PRT; 228 AA.
AC Q9AK63;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Putative integral membrane protein.
GN SCO4054 OR 2SCD60.20.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AB939118; CAC32325.1; -
KW Complete proteome.
SQ SEQUENCE 228 AA; 25240 MW; C13D8696B40D2991 CRC64;

Query Match 52.9%; Score 46; DB 16; Length 228;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MLPULLGLGPACW 15
Db 1 MLPULLGLGPACW 15

RESULT 12
Q967F2 PRELIMINARY; PRT; 496 AA.
AC Q967F2;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical 57.2 kDa protein (Putative alpha 2
DE mannosyltransferase).
GN T27F7.3 OR PIG-B.

OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Waterston R.;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Hallsworth K.;
RT "The sequence of C. elegans cosmid T27F7.";
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.

RA Oriol R., Martinez-Duncker I., Chantret I., Mollicone R., Codogno P.;
RT "Common origin and evolution of glycosyltransferases using Dol-P-
RT monosaccharides as donor substrate.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; U58762; AAK39304.2; -
DR EMBL; AJ431373; CAD24083.1; -
DR WormPep; T27F7.3a; CE30458.
DR InterPro; IPR005599; PMP.
DR Pfam; PF03901; PMP; 1.
KW Hypothetical protein; Glycosyltransferase; Transferase.
SQ SEQUENCE 496 AA; 57191 MW; 6A0C53E5346D7F8C CRC64;

Query Match 52.9%; Score 46; DB 5; Length 496;
Best Local Similarity 63.6%; Pred. No. 59;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPULLGLGP 11
Db 270 MLPILLGLGP 280

RESULT 13
Q8CHQ8 PRELIMINARY; PRT; 517 AA.
AC Q8CHQ8;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Similar to N-acetylglucosamine-1-phosphodiester
DE alpha-N-acetylglucosaminidase.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N;
RA Strausberg R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC039790; AAH39790.1; -
SQ SEQUENCE 517 AA; 56013 MW; 80F6E9AFB3873177 CRC64;

Query Match 52.9%; Score 46; DB 11; Length 517;
Best Local Similarity 64.3%; Pred. No. 61;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLPULLGLGPAC 14
Db 1 MLPULLGLGPAC 14

Db 11 LIPALLGLGVAWC 24

RESULT 14

Q8BJ48 PRELIMINARY; PRT; 517 AA.
ID Q8BJ48
AC Q8BJ48;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE N-acetylglucosamine-1-PHOSPHODIESTER
DE alpha-N-acetylglucosaminidase.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK032158; BAC27731.1;
SQ SEQUENCE 517 AA; 56056 MW; 177F813D20552C4B CRC64;

Query Match 52.9%; Score 46; DB 11; Length 517;
Best Local Similarity 64.3%; Pred. No. 61;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLPLLGLGPAC 14
Db 11 LIPALLGLGVAWC 24

RESULT 15

Q9UVMS PRELIMINARY; PRT; 558 AA.
ID Q9UVMS
AC Q9UVMS;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Mfs1.2.
GN MFS1.2.
OS Coprinus cinereus (inky cap fungus).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Psathyrellaceae; Coprinopsis.
OX NCBI_TaxID=5346;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JV6;
RX MEDLINE=20221720; PubMed=10757757;
RA Halsall J.R., Milner M.J., Casselton L.A.;
RT "Three Subfamilies of Pheromone and Receptor Genes Generate Multiple B
RT Mating Specificities in the Mushroom Coprinus cinereus.";
RL Genetics 154:1115-1123(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AF186392; AAF01427.1; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR001411; TCR_TetB.
DR Pfam; PF00083; sugar_tr_1.
DR PRINTS; PR01036; TCR_TetB.
DR PROSITE; PS50850; MFS; 1.
KW Transmembrane.
SQ SEQUENCE 558 AA; 59446 MW; 6F3232638C26A2AF CRC64;

Query Match 52.9%; Score 46; DB 3; Length 558;
Best Local Similarity 53.3%; Pred. No. 66;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLPLLGLGPACW 15
Db 255 LVPLILGVLIAWM 269

Search completed: January 12, 2004, 08:19:14
Job time : 7.20359 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:09 ; Search time 2.41118 Seconds
(without alignments)
1053.272 Million cell updates/sec

Title: US-10-085-167-2_COPY_1_16
Perfect score: 87

Sequence: 1 MLPLLLGLGPACWA 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

1: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-----------------------------|
| 1 | 87 | 100.0 | 17 | 22 | AAB61467 MANGO 245 signal p |
| 2 | 87 | 100.0 | 133 | 23 | ABG70386 Adipocyte compleme |
| 3 | 87 | 100.0 | 284 | 23 | ABG70384 Adipocyte compleme |
| 4 | 87 | 100.0 | 299 | 23 | ABG70385 Adipocyte compleme |
| 5 | 87 | 100.0 | 329 | 22 | AAB61424 Monkey MANGO 245 p |
| 6 | 87 | 100.0 | 329 | 22 | AAB61606 Human ZACRP4. Hom |
| 7 | 87 | 100.0 | 329 | 23 | ABG79643 Human novel secret |
| 8 | 87 | 100.0 | 348 | 22 | AAB61423 Human MANGO 245 pr |
| 9 | 73 | 83.9 | 199 | 22 | AAB61488 Murine MANGO 245 p |

| | | | | | |
|----|------|------|------|----|-----------------------------|
| 10 | 72 | 82.8 | 14 | 22 | AAB61465 MANGO 245 signal p |
| 11 | 52 | 59.8 | 115 | 21 | AAV52384 Pseudomonas aerugi |
| 12 | 49 | 56.3 | 1061 | 22 | ABB65553 Drosophila melanog |
| 13 | 49 | 56.3 | 1180 | 19 | AAW61273 Homo sapiens proto |
| 14 | 48 | 55.2 | 912 | 22 | AAU52873 Propionibacterium |
| 15 | 47 | 54.0 | 167 | 22 | ABG17123 Novel human diagno |
| 16 | 47 | 54.0 | 516 | 23 | ABB06122 Human NS protein s |
| 17 | 47 | 54.0 | 656 | 19 | AAW71648 Fluorescent calmod |
| 18 | 45.5 | 52.3 | 101 | 22 | AAO10317 Human polypeptide |
| 19 | 45.5 | 52.3 | 120 | 22 | AAO10647 Human polypeptide |
| 20 | 45 | 51.7 | 16 | 21 | AAV93254 Signal peptide use |
| 21 | 45 | 51.7 | 18 | 19 | AAW71681 Endoplasmic reticu |
| 22 | 45 | 51.7 | 18 | 22 | AAB98994 N-terminal calreti |
| 23 | 45 | 51.7 | 92 | 22 | ABB16457 Human nervous syst |
| 24 | 45 | 51.7 | 347 | 21 | AAB42695 Human ORFX ORF2459 |
| 25 | 45 | 51.7 | 378 | 24 | ABU07370 Human protein NOV2 |
| 26 | 45 | 51.7 | 602 | 22 | ABB71195 Drosophila melanog |
| 27 | 44 | 50.6 | 54 | 22 | ABG08743 Novel human diagno |
| 28 | 44 | 50.6 | 90 | 22 | AAM84788 Human immune/haema |
| 29 | 44 | 50.6 | 318 | 17 | AAR88105 Bovine p32 11-cis- |
| 30 | 44 | 50.6 | 318 | 18 | AAW18343 Human protein sequ |
| 31 | 44 | 50.6 | 390 | 22 | AAB92502 Human PRO731 prote |
| 32 | 44 | 50.6 | 1184 | 20 | AAV41750 Human PRO731 (UNQ3 |
| 33 | 44 | 50.6 | 1184 | 21 | AAB44306 Human PRO731 polyp |
| 34 | 44 | 50.6 | 1184 | 22 | AAU12368 Human PRO731 polyp |
| 35 | 44 | 50.6 | 1184 | 24 | ABU66766 Human PRO polypept |
| 36 | 44 | 50.6 | 1184 | 24 | ABU67042 Human secreted/cra |
| 37 | 44 | 50.6 | 1184 | 24 | ABU59847 Novel secreted and |
| 38 | 44 | 50.6 | 1184 | 24 | ABU61136 Human PRO731 polyp |
| 39 | 44 | 50.6 | 1184 | 24 | ABU03542 Angiogenesis-assoc |
| 40 | 44 | 50.6 | 1187 | 22 | AAE06028 Human asthma-assoc |
| 41 | 44 | 50.6 | 1187 | 22 | AAB95684 Human protein sequ |
| 42 | 43 | 49.4 | 76 | 22 | AAW06512 Human foetal prote |
| 43 | 43 | 49.4 | 96 | 17 | AAW00723 Murine fibrosin 28 |
| 44 | 43 | 49.4 | 96 | 17 | AAW11890 Murine fibrosin 28 |
| 45 | 43 | 49.4 | 119 | 21 | AAV75262 Neisseria gonorrhe |

ALIGNMENTS

| | |
|---|--|
| RESULT 1 | |
| AAB61467 | |
| ID AAB61467 standard; peptide; 17 AA. | |
| XX | |
| AC AAB61467; | |
| XX | |
| DT 04-APR-2001 (first entry) | |
| XX | |
| DE MANGO 245 signal peptide #2. | |
| XX | |
| KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung; | |
| KW autoimmune; allergy; cardiovascular; brain; degenerative; placental; | |
| KW pancreatic; skeletal; muscle. | |
| XX | |
| OS Catarrhini sp.. | |
| XX | |
| PN WO200100672-A1. | |
| XX | |
| PD 04-JAN-2001. | |
| XX | |
| PF 29-JUN-2000; 2000WO-US18184. | |
| XX | |
| PR 29-JUN-1999; 99US-0342687. | |
| XX | |
| PA (MILL-) MILLENNIUM PHARM INC. | |
| XX | |
| PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD; | |
| XX | |
| DR WPI; 2001-050127/06. | |
| XX | |
| PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and | |
| PT MANGO 245 proteins, useful in the treatment of inflammatory diseases | |

| | |
|----------|--|
| PT | (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver disorders (e.g. jaundice) - |
| XX | |
| PS | Disclosure; Page 243; 262pp; English. |
| XX | |
| CC | The present invention relates to cDNAs encoding TANGO 244, |
| CC | TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins. |
| CC | The nucleic acids, proteins and protein modulators are useful for treating colonic disorders, inflammatory diseases, tumors, renal disorders, liver disorders, lung disorders, autoimmune diseases, allergic diseases, cardiovascular diseases, brain disorders, degenerative diseases placental, pancreatic, skeletal and muscle disorders. |
| CC | |
| XX | |
| SQ | Sequence 17 AA; |
| | |
| | Query Match 100.0%; Score 87; DB 22; Length 17; |
| | Best Local Similarity 100.0%; Pred. No. 3.2e-07; |
| | Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |
| OY | 1 MLPLLLGLGPACWA 16 1 MLPLLLGLGPACWA 16 |
| Dn | |
| RESULT 2 | |
| ABG70386 | ABG70386 standard; Protein; 133 AA. |
| ID | |
| XX | |
| AC | ABG70386; |
| XX | |
| DT | 05-NOV-2002 (first entry) |
| XX | |
| DE | Adipocyte complement-related C1q Tumour Necrosis Factor-like protein #4. |
| XX | |
| KM | Human; NOX; NOX-associated disorder; cardiomyopathy; atherosclerosis; cell signal processing; metabolic pathway modulation; metabolic disorder; obesity; diabetes; infectious disease; neurodegenerative disorder; acne; Alzheimer's disease; Parkinson's disease; immune disorder; cancer; haematopoietic disorder; cirrhosis; pancreatitis; learning defect; memory defect; infertility; congenital heart defect; hair growth; pigmentation disorder; endocrine disorder; respiratory disease; health; gastro-intestinal disease; reproductive; neurological disease; bone marrow transplantation; endocrine disease; allergy; inflammation; nephrological disorder; urinary system disorder; age-related disorder; neuropsychiatric disorder; EGF-related protein; SCUB1; TEN-M4; adipocyte complement-related C1q tumour necrosis factor; out at first; beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter; type Ia membrane sushi domain containing; butyrophilin; type Ia membrane-sushi domain containing. |
| KW | |
| KW | |
| OS | Homo sapiens. |
| XX | |
| PN | WO200257453-A2. |
| XX | |
| PD | 25-JUL-2002. |
| XX | |
| PF | 19-DEC-2001; 2001WO-US50331. |
| XX | |
| PR | 19-DEC-2000; 2000US-265704P. |
| XX | |
| PR | 20-DEC-2000; 2000US-257314P. |
| PR | 02-MAY-2001; 2001US-288153P. |
| PR | 29-MAY-2001; 2001US-294075P. |
| PR | 24-JUL-2001; 2001US-307506P. |
| PR | 10-AUG-2001; 2001US-311590P. |
| PR | 10-AUG-2001; 2001US-311613P. |
| PR | 29-AUG-2001; 2001US-315617P. |
| PR | 14-SEP-2001; 2001US-322358P. |
| XX | |
| PA | (CURA-) CURAGEN CORP. |
| XX | |
| PI | Gangoli EA, Patturajan M, Vernet CAM, Maliyankar UM, Kekuda R, Stone DJ, Anderson D, Shinkets RA, Burgess CE, Zerhusen BD, Liu X; |

| | | | | | | |
|----------|---|---|------------|-------------|-------|-------|
| PI | Spytek KA, | Casman SJ, | Boldog FL, | Smithson G, | Li L, | Ji W; |
| XX | | | | | | |
| DR | WPI; 2002-590744/63. | | | | | |
| DR | N-PSDB; ABS52098. | | | | | |
| XX | | | | | | |
| PT | Novel isolated NOVX polypeptide useful for treating cardiomyopathy, | | | | | |
| PT | atherosclerosis, metabolic disorders, diabetes, obesity, infectious | | | | | |
| PT | disease, anorexia, neurodegenerative disorders, Alzheimer's disease or | | | | | |
| PT | cancer - | | | | | |
| XX | | | | | | |
| PS | Claim 1; Page 32; 318pp; English. | | | | | |
| XX | | | | | | |
| CC | The present invention relates to new NOVX polypeptides. The invention is | | | | | |
| CC | useful for treating or preventing a NOVX-associated disorder such as | | | | | |
| CC | cardiomyopathy or atherosclerosis, where the disorder is related to cell | | | | | |
| CC | signal processing and metabolic pathway modulation in a subject, | | | | | |
| CC | preferably human. The invention is also useful for treating metabolic | | | | | |
| CC | disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative | | | | | |
| CC | disorders (e.g. Alzheimer's disease, Parkinson's disease), immune | | | | | |
| CC | disorders, haematopoietic disorders and various cancers. The molecules of | | | | | |
| CC | the invention are also useful for treating or preventing cirrhosis, | | | | | |
| CC | pancreatitis, learning and memory defects, infertility, congenital heart | | | | | |
| CC | defects, acne, hair growth, pigmentation disorders, endocrine disorders, | | | | | |
| CC | respiratory disease, gastro-intestinal diseases, reproductive, health, | | | | | |
| CC | neurological diseases, bone marrow transplantation, endocrine diseases, | | | | | |
| CC | allergy and inflammation, nephrological disorders, urinary system | | | | | |
| CC | disorders, neuropsychiatric disorders and age-related disorders. | | | | | |
| CC | The present amino acid sequence represents a NOVX protein of the | | | | | |
| CC | invention. | | | | | |
| XX | | | | | | |
| SQ | Sequence | 133 AA; | | | | |
| | Query Match | 100.0%; Score 87; DB 23; Length 133; | | | | |
| | Best Local Similarity | 100.0%; Pred. No. 2.9e-06; | | | | |
| | Matches | 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | |
| QY | | | | | | |
| | 1 MLPULLGLGPACWA 16 | | | | | |
| | | | | | | |
| | 1 MLPULLGLGPACWA 16 | | | | | |
| Db | | | | | | |
| | | | | | | |
| RESULT 3 | | | | | | |
| ABG70384 | | | | | | |
| ID | ABG70384 standard; Protein; 284 AA. | | | | | |
| XX | | | | | | |
| AC | ABG70384; | | | | | |
| XX | | | | | | |
| DT | 05-NOV-2002 (first entry) | | | | | |
| XX | | | | | | |
| DE | Adipocyte complement-related C1q Tumour Necrosis Factor-like protein #2. | | | | | |
| XX | | | | | | |
| KM | Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis; | | | | | |
| KM | cell signal processing; metabolic pathway modulation; metabolic disorder; | | | | | |
| KM | obesity; diabetes; infectious disease; neurodegenerative disorder; acne; | | | | | |
| KM | Alzheimer's disease; Parkinson's disease; immune disorder; cancer; | | | | | |
| KM | haematopoietic disorder; cirrhosis; pancreatitis; learning defect; | | | | | |
| KM | memory defect; infertility; congenital heart defect; hair growth; | | | | | |
| KM | pigmentation disorder; endocrine disorder; respiratory disease; health; | | | | | |
| KM | gastro-intestinal disease; reproductive; neurological disease; | | | | | |
| KM | bone marrow transplantation; endocrine disease; allergy; inflammation; | | | | | |
| KM | nephrological disorder; urinary system disorder; age-related disorder; | | | | | |
| KM | neuropsychiatric disorder; EGF-related protein; SCUBE1; TEN-M4; | | | | | |
| KM | adipocyte complement-related C1q tumour necrosis factor; out at first; | | | | | |
| KM | beta adrenergic receptor kinase; EphA6/enk-2; glucose transporter; | | | | | |
| KM | type Ia membrane sushi-containing domain; butyrophilin; | | | | | |
| KM | type Ia membrane-sushi domain containing. | | | | | |
| OS | Homo sapiens. | | | | | |
| XX | | | | | | |
| PN | WO200257453-A2. | | | | | |
| XX | | | | | | |
| PD | 25-JUL-2002. | | | | | |
| XX | | | | | | |

PF 19-DEC-2001; 2001WO-US50331.
XX
PR 19-DEC-2000; 2000US-265704P.
PR 20-DEC-2000; 2000US-257314P.
PR 02-MAY-2001; 2001US-288153P.
PR 29-MAY-2001; 2001US-294075P.
PR 24-JUL-2001; 2001US-307506P.
PR 10-AUG-2001; 2001US-311590P.
PR 10-AUG-2001; 2001US-311613P.
PR 29-AUG-2001; 2001US-315617P.
PR 14-SEP-2001; 2001US-322358P.
XX
XX
PA (CURA-) CURAGEN CORP.
XX
PI Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
PI Stone DJ, Anderson D, Shimkets RA, Burgess CE, Zerhusen BD, Liu X;
PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
DR MPI; 2002-590744/63.
DR N-PSDB; ABS52096.
XX
XX
PT Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious
PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
PT cancer -
XX
XX
PS Claim 1; Page 30; 318bp; English.
XX
XX
CC The present invention relates to new NOVX polypeptides. The invention is
CC useful for treating or preventing a NOVX-associated disorder such as
CC cardiomyopathy or atherosclerosis, where the disorder is related to cell
CC signal processing and metabolic pathway modulation in a subject,
CC preferably human. The invention is also useful for treating metabolic
CC disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
CC disorders, haematopoietic disorders and various cancers. The molecules of
CC the invention are also useful for treating or preventing cirrhosis,
CC pancreatitis, learning and memory defects, infertility, congenital heart
CC defects, acne, hair growth, pigmentation disorders, endocrine disorders,
CC respiratory disease, gastro-intestinal diseases, reproductive, health,
CC neurological diseases, bone marrow transplantation, endocrine diseases,
CC allergy and inflammation, nephrological disorders, urinary system
CC disorders, neuropsychiatric disorders and age-related disorders.
CC The present amino acid sequence represents a NOVX protein of the
CC invention.
XX
XX
SQ Sequence 284 AA;
XX
XX
Query Match 100.0%; Score 87; DB 23; Length 284;
Best Local Similarity 100.0%; Pred. No. 6.7e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MLPLLLGLGPACWA 16
Db 1 MLPLLLGLGPACWA 16
XX
XX
RESULT 4
ABG70385
ID ABG70385 standard; Protein; 299 AA.
XX
XX
AC ABG70385;
XX
XX
DT 05-NOV-2002 (first entry)
XX
DE Adipocyte complement-related C1q Tumour Necrosis Factor-like protein #3.
XX
XX
KW Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis;
KW cell signal processing; metabolic pathway modulation; metabolic disorder;
KW obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
KW Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
KW haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
KW memory defect; infertility; congenital heart defect; hair growth;

KW pigmentation disorder; endocrine disorder; respiratory disease; health;
KW gastro-intestinal disease; reproductive; neurological disease;
KW bone marrow transplantation; endocrine disease; allergy; inflammation;
KW nephrological disorder; urinary system disorder; age-related disorder;
KW neuropsychiatric disorder; EGF-related protein; SCUB1; TEN-M4;
KW adipocyte complement-related C1q tumour necrosis factor; out at first;
KW beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;
KW type 1a membrane sushi-containing domain; butyrophilin;
XX
XX
OS Homo sapiens.
XX
XX
PN WO200257453-A2.
XX
XX
PD 25-JUL-2002.
XX
XX
PF 19-DEC-2001; 2001WO-US50331.
XX
XX
PR 19-DEC-2000; 2000US-265704P.
PR 20-DEC-2000; 2000US-257314P.
PR 02-MAY-2001; 2001US-288153P.
PR 29-MAY-2001; 2001US-294075P.
PR 24-JUL-2001; 2001US-307506P.
PR 10-AUG-2001; 2001US-311590P.
PR 10-AUG-2001; 2001US-311613P.
PR 29-AUG-2001; 2001US-315617P.
PR 14-SEP-2001; 2001US-322358P.
XX
XX
PA (CURA-) CURAGEN CORP.
XX
XX
PI Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
PI Stone DJ, Anderson D, Shimkets RA, Burgess CE, Zerhusen BD, Liu X;
PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
XX
DR MPI; 2002-590744/63.
DR N-PSDB; ABS52097.
XX
XX
PT Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious
PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
PT cancer -
XX
XX
PS Claim 1; Page 31; 318bp; English.
XX
XX
CC The present invention relates to new NOVX polypeptides. The invention is
CC useful for treating or preventing a NOVX-associated disorder such as
CC cardiomyopathy or atherosclerosis, where the disorder is related to cell
CC signal processing and metabolic pathway modulation in a subject,
CC preferably human. The invention is also useful for treating metabolic
CC disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
CC disorders, haematopoietic disorders and various cancers. The molecules of
CC the invention are also useful for treating or preventing cirrhosis,
CC pancreatitis, learning and memory defects, infertility, congenital heart
CC defects, acne, hair growth, pigmentation disorders, endocrine disorders,
CC respiratory disease, gastro-intestinal diseases, reproductive, health,
CC neurological diseases, bone marrow transplantation, endocrine diseases,
CC allergy and inflammation, nephrological disorders, urinary system
CC disorders, neuropsychiatric disorders and age-related disorders.
CC The present amino acid sequence represents a NOVX protein of the
CC invention.
XX
XX
SQ Sequence 299 AA;
XX
XX
Query Match 100.0%; Score 87; DB 23; Length 299;
Best Local Similarity 100.0%; Pred. No. 7e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MLPLLLGLGPACWA 16
Db 1 MLPLLLGLGPACWA 16

```
RESULT 5
AAB61424
ID AAB61424 standard; protein; 329 AA.
XX
AC AAB61424;
XX
DT 04-APR-2001 (first entry)
XX
DE Monkey MANGO 245. protein.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Catarrhini sp..
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
PS WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Claim 1; Fig 25; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 329 AA;
Query Match 100.0%; Score 87; DB 22; Length 329;
Best Local Similarity 100.0%; Pred. No. 7.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MLPLLLGLGPACWA 16
Db 1 MLPLLLGLGPACWA 16
RESULT 6
AAB61606
ID AAB61606 standard; Protein; 329 AA.
XX
AC AAB61606;
XX
DT 05-APR-2001 (first entry)
XX
DE Human ZACRP4.
XX
KW Human; zacrpa; complement factor C1q domain; chromosome 11q11;
KW energy balance; cellular metabolic reaction; autocrine factor;
KW development; cell proliferation; differentiation; cell survival.
XX
OS Homo sapiens.
XX
```

```
FH Key Location/Qualifiers
FT Peptide 1..16
FT /label= Signal_peptide
FT Domain 17..159
FT /label= C1q_domain_#1
FT Domain 160..328
FT /label= C1q_domain_#2
XX
PN WO200102565-A2.
XX
PD 11-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US17692.
XX
PR 01-JUL-1999; 99US-0346502.
XX
PA (ZYMO ) ZYMOGENETICS INC.
XX
PI Holloway JL, Lok S;
XX
DR WPI; 2001-138140/14.
DR N-PSDB; AAF28672.
XX
PT Novel secreted protein ZACRP4 polypeptides having tandem C1q globular
PT domains, useful for studying cell-cell communication and regulation of
PT cellular processes -
XX
PS Claim 1; Page 77-78; 82pp; English.
XX
CC The present sequence is human ZACRP4 protein. ZACRP4 protein has two
CC complement factor C1q domains. The ZACRP4 gene is located on human
CC chromosome 11q11. The ZACRP4 coding sequence and protein have a number of
CC uses described in the specification, including, modulation of energy
CC balance and cellular metabolic reactions in mammals. In addition, ZACRP4
CC protein is useful as an autocrine factor, particularly during
CC development, in mediating the processes of an organism, in regulating
CC cellular processes such as cell proliferation and/or differentiation,
CC cell survival and energy balance.
XX
SQ Sequence 329 AA;
Query Match 100.0%; Score 87; DB 22; Length 329;
Best Local Similarity 100.0%; Pred. No. 7.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MLPLLLGLGPACWA 16
Db 1 MLPLLLGLGPACWA 16
RESULT 7
ABG79643
ID ABG79643 standard; Protein; 329 AA.
XX
AC ABG79643;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human novel secreted protein SECP19, Incyte ID No. 931619CD1.
XX
KW Human; SECP; secreted protein; cell proliferative disorder;
KW actinic keratosis; arteriosclerosis; bursitis; hepatitis; cancer;
KW autoimmune disorder; inflammatory disorder; AIDS; asthma; allergy;
KW acquired immunodeficiency syndrome; anaemia; atopic dermatitis;
KW cardiovascular disorder; congestive heart failure; vascular tumour;
KW ischaemic heart disease; myocardial infarction; epilepsy; stroke;
KW Alzheimer's disease; neurological disorder; cerebral neoplasia;
KW Cushing's syndrome; Duchenne muscular dystrophy; hypothyroidism;
KW Becker muscular dystrophy.
XX
OS Homo sapiens.
XX
```

PN WO200262841-A2.
XX
PD 15-AUG-2002.
XX
PF 28-JAN-2002; 2002WO-US02616.
XX
PR 02-FEB-2001; 2001US-266195P.
PR 08-FEB-2001; 2001US-267924P.
PR 09-FEB-2001; 2001US-267816P.
PR 09-FEB-2001; 2001US-268112P.
PR 26-FEB-2001; 2001US-271639P.
PR 07-SEP-2001; 2001US-317818P.
PR 21-DEC-2001; 2001US-343553P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Tang TY, Yue H, Gandhi AR, Yao MG, Warren BA, Ding L, Duggan BM;
PI Xu Y, Yang J, Thangavelu K, Lal PG, Honchell CD, Walia NK, Lee S;
PI Lee EA, Richardson TW, Baughn MR, Elliott VS;
XX
XX WPI; 2002-657522/70.
DR N-PSDB; ABS64954.
XX
XX
PT New human secreted proteins and nucleic acids useful in diagnosing,
PT treating and preventing cell proliferative, autoimmune/inflammatory,
PT cardiovascular, neurological, and developmental disorders -
XX
XX
PS Claim 1; Page 140; 158pp; English.
XX
XX The invention relates to twenty four human secreted proteins
CC (SECP1-24), proteins 90% identical to them and active fragments of them.
CC Also included are nucleic acids encoding the SECP proteins, a recombinant
CC polynucleotide comprising a promoter sequence operably linked to the
CC nucleic acid, a cell transformed with the recombinant polynucleotide,
CC a transgenic organism comprising the recombinant polynucleotide, an
CC anti-SECP antibody, and screening for ant/agonists and modulators of
CC SECP function or expression. The SECP proteins and nucleic acids are
CC useful in the diagnosis, treatment and prevention of cell proliferative
CC (e.g. actinic keratosis, arteriosclerosis, bursitis, hepatitis or
CC cancer), autoimmune/inflammatory (e.g. AIDS (acquired immunodeficiency
CC syndrome), asthma, anaemia, allergies or atopic dermatitis),
CC cardiovascular (e.g. congestive heart failure, ischaemic heart disease,
CC myocardial infarction, hypertensive heart disease, or vascular tumours),
CC neurological (e.g. epilepsy, stroke, cerebral neoplasms, or Alzheimer's
CC disease), and developmental (e.g. renal tubular acidosis, Cushing's
CC syndrome, Duchenne and Becker muscular dystrophy, or hypothyroidism)
CC disorders. Many other diseases and disorders are listed in the
CC specification. These may also be used in assessing the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of the secreted proteins. The present sequence represents a
CC SECP protein of the invention.
XX
XX
SQ Sequence 329 AA;
XX
XX
Query Match 100.0%; Score 87; DB 23; Length 329;
Best Local Similarity 100.0%; Pred. No. 7.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MLPLLLGLGPACWA 16
Db 1 MLPLLLGLGPACWA 16
XX
XX
RESULT 8
AAB61423
ID AAB61423 standard; protein; 348 AA.
XX
XX
AC AAB61423;
XX
DT 04-APR-2001 (first entry)
XX
DE Human MANGO 245 protein.
XX

KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
XX Homo sapiens.
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
XX WPI; 2001-050127/06.
DR
XX
XX Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
XX
PS Claim 1; Fig 23; 262pp; English.
XX
XX The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
XX
SQ Sequence 348 AA;
XX
XX
Query Match 100.0%; Score 87; DB 22; Length 348;
Best Local Similarity 100.0%; Pred. No. 8.3e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MLPLLLGLGPACWA 16
Db 1 MLPLLLGLGPACWA 16
XX
XX
RESULT 9
AAB61488
ID AAB61488 standard; protein; 199 AA.
XX
XX
AC AAB61488;
XX
DT 04-APR-2001 (first entry)
XX
DE Murine MANGO 245 protein.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
XX Mus musculus.
XX
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX

PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Claim 1; Fig 29; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 199 AA;
XX
Query Match 83.9%; Score 73; DB 22; Length 199;
Best Local Similarity 87.5%; Pred. No. 0.00082;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MLPLLLGLGPACWA 16
|| ||||| |||||
Db 1 MLPLLLGLGPACWA 16
RESULT 10
AAB61465
ID AAB61465 standard; peptide; 14 AA.
XX
AC AAB61465;
XX
DT 04-APR-2001 (first entry)
XX
DE MANGO 245 signal peptide #1.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Homo sapiens.
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Disclosure; Page 242; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,

CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 14 AA;
XX
Query Match 82.8%; Score 72; DB 22; Length 14;
Best Local Similarity 100.0%; Pred. No. 6.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLPLLLGLGPAC 14
|| ||||| |||||
Db 1 MLPLLLGLGPAC 14
RESULT 11
AAY52384
ID AAY52384 standard; Protein; 115 AA.
XX
AC AAY52384;
XX
DT 09-FEB-2000 (first entry)
XX
DE Pseudomonas aeruginosa PilY2 protein.
XX
KW Pilus; motility; biofilm; biotic surface; abiotic surface;
KW nutritional status; environment; gene expression; cellular physiology;
KW antibiotic resistance; gene product; transcription;
KW control; manufacturing; agriculture; healthcare; identification;
KW screening; modulation; formation; colonisation; plant root;
KW contact lens; medical implant; catheter wall; sewerage pipe;
KW water pipe; lung infection.
XX
OS Pseudomonas aeruginosa.
XX
PN WO9955368-A1.
XX
PD 04-NOV-1999.
XX
PF 27-APR-1999; 99WO-US09034.
XX
PR 27-APR-1998; 98US-0083259.
PR 02-OCT-1998; 98US-0102870.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI O'Toole GA, Kolter R;
XX
DR WPI; 2000-023264/02.
XX
PT Isolated surface attachment defective genes, useful for developing
PT products which modulate bacterial biofilm formation -
XX
PS Claim 32; Page 149; 151pp; English.
XX
CC Sequences AAY52377-Y52385 are involved in pilus synthesis and motility
CC and are required for biofilm formation by Pseudomonas aeruginosa. The
CC formation of bacterial biofilms occurs when bacteria make the transition
CC from a planktonic (free swimming) existence to a lifestyle in which the
CC microorganisms are firmly attached to biotic or abiotic surfaces. This
CC transition is thought to be regulated in part by the nutritional status
CC of the environment and is accompanied by a change in gene expression
CC patterns, cellular physiology and higher resistance to antibiotics
CC relative to planktonic bacterial counterparts. Pilus synthetic and
CC motility genes, gene products, and transcriptional control regions may
CC all be used for the control of biofilm formation in fields such as
CC manufacturing, agriculture, and healthcare. The products may also be
CC used to identify compounds which modulate biofilm formation. Compounds
CC that stimulate biofilm formation can be used to improve colonisation of
CC plant roots by beneficial bacteria (e.g., nitrogen fixing bacteria).
CC Conversely, compounds that inhibit biofilm formation can be used to
CC restrict the growth of bacteria on contact lenses, medical implants
CC (e.g., artificial hips), the walls of catheters, fluid contacting

AC AAU52873;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #13769.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KM dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
DR N-PSDB; AAS59557.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
XX Example 1; SEQ ID No 14068; 10699p; English.
PS
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 912 AA;
Query Match 55.2%; Score 48; DB 22; Length 912;
Best Local Similarity 69.2%; Pred. No. 45;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 4 LLLGLLPACWA 16
|||:|:|
Db 373 LILGLTGPSSGWA 385

RESULT 15
ABG17123
ID ABG17123 standard; Protein; 167 AA.
XX

AC ABG17123;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #17114.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR N-PSDB; AAS81310.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 20; SEQ ID No 47482; 103p; English.
PS
XX

The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX

SQ Sequence 167 AA;
Query Match 54.0%; Score 47; DB 22; Length 167;
Best Local Similarity 58.8%; Pred. No. 10;
Matches 10; Conservative 2; Mismatches 3; Indels 2; Gaps 1;
OY 2 LPLLL-GLLPACWA 16
|||:|:|
Db 68 LPLLVLTGLFGMAECWS 84

Search completed: January 12, 2004, 08:15:38
Job time : 4.41118 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:15:44 ; Search time 1.9481 Seconds
(without alignments)
1655.023 Million cell updates/sec

Title: US-10-085-167-2_COPY_1_16
Perfect score: 87
Sequence: 1 MLPLLLGLGPACWA 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747907 seqs, 201509753 residues

Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|---------------------|-------------------|
| 1 | 87 | 100.0 | 329 | 12 | US-10-236-055A-14 | Sequence 14, Appl |
| 2 | 87 | 100.0 | 329 | 12 | US-10-085-167-2 | Sequence 2, Appl |
| 3 | 73 | 83.9 | 326 | 12 | US-10-236-055A-16 | Sequence 16, Appl |
| 4 | 49 | 56.3 | 1135 | 11 | US-09-759-130B-42 | Sequence 42, Appl |
| 5 | 48 | 55.2 | 379 | 12 | US-10-120-801-40 | Sequence 40, Appl |
| 6 | 47 | 54.0 | 656 | 10 | US-09-554-000-8 | Sequence 8, Appl |
| 7 | 45 | 51.7 | 18 | 10 | US-09-554-000-56 | Sequence 56, Appl |
| 8 | 45 | 51.7 | 347 | 12 | US-10-120-801-41 | Sequence 41, Appl |
| 9 | 45 | 51.7 | 378 | 12 | US-10-120-801-4 | Sequence 4, Appl |
| 10 | 45 | 51.7 | 416 | 12 | US-10-316-253-2 | Sequence 2, Appl |
| 11 | 45 | 51.7 | 416 | 12 | US-10-316-253-4 | Sequence 6, Appl |
| 12 | 45 | 51.7 | 416 | 12 | US-10-316-253-6 | Sequence 4, Appl |
| 13 | 44 | 50.6 | 242 | 12 | US-10-369-493-4780 | Sequence 4780, Ap |
| 14 | 44 | 50.6 | 242 | 12 | US-10-369-493-7540 | Sequence 7540, Ap |
| 15 | 44 | 50.6 | 394 | 15 | US-10-156-761-10413 | Sequence 10413, A |

| | | | | | | |
|----|----|------|------|----|--------------------|-------------------|
| 16 | 44 | 50.6 | 615 | 10 | US-09-908-193-26 | Sequence 26, Appl |
| 17 | 44 | 50.6 | 615 | 10 | US-09-908-193-27 | Sequence 27, Appl |
| 18 | 44 | 50.6 | 1184 | 10 | US-09-978-295A-425 | Sequence 425, App |
| 19 | 44 | 50.6 | 1184 | 10 | US-09-978-697-425 | Sequence 425, App |
| 20 | 44 | 50.6 | 1184 | 10 | US-09-978-192A-425 | Sequence 425, App |
| 21 | 44 | 50.6 | 1184 | 10 | US-09-999-832A-425 | Sequence 425, App |
| 22 | 44 | 50.6 | 1184 | 11 | US-09-978-189-425 | Sequence 425, App |
| 23 | 44 | 50.6 | 1184 | 11 | US-09-978-608A-425 | Sequence 425, App |
| 24 | 44 | 50.6 | 1184 | 11 | US-09-978-585A-425 | Sequence 425, App |
| 25 | 44 | 50.6 | 1184 | 11 | US-09-978-191A-425 | Sequence 425, App |
| 26 | 44 | 50.6 | 1184 | 11 | US-09-978-403A-425 | Sequence 425, App |
| 27 | 44 | 50.6 | 1184 | 11 | US-09-978-564A-425 | Sequence 425, App |
| 28 | 44 | 50.6 | 1184 | 11 | US-09-999-833A-425 | Sequence 425, App |
| 29 | 44 | 50.6 | 1184 | 11 | US-09-981-915A-425 | Sequence 425, App |
| 30 | 44 | 50.6 | 1184 | 11 | US-09-978-824-425 | Sequence 425, App |
| 31 | 44 | 50.6 | 1184 | 11 | US-09-918-585A-425 | Sequence 425, App |
| 32 | 44 | 50.6 | 1184 | 11 | US-09-978-423A-425 | Sequence 425, App |
| 33 | 44 | 50.6 | 1184 | 11 | US-09-978-193A-425 | Sequence 425, App |
| 34 | 44 | 50.6 | 1184 | 11 | US-09-999-830A-425 | Sequence 425, App |
| 35 | 44 | 50.6 | 1184 | 11 | US-09-978-757A-425 | Sequence 425, App |
| 36 | 44 | 50.6 | 1184 | 11 | US-09-978-187B-425 | Sequence 425, App |
| 37 | 44 | 50.6 | 1184 | 11 | US-09-978-643A-425 | Sequence 425, App |
| 38 | 44 | 50.6 | 1184 | 12 | US-09-978-375A-425 | Sequence 425, App |
| 39 | 44 | 50.6 | 1184 | 12 | US-09-978-188A-425 | Sequence 425, App |
| 40 | 44 | 50.6 | 1184 | 12 | US-09-978-298A-425 | Sequence 425, App |
| 41 | 44 | 50.6 | 1184 | 12 | US-10-137-870-394 | Sequence 394, App |
| 42 | 44 | 50.6 | 1184 | 12 | US-10-140-018-394 | Sequence 394, App |
| 43 | 44 | 50.6 | 1184 | 12 | US-10-140-021-394 | Sequence 394, App |
| 44 | 44 | 50.6 | 1184 | 12 | US-10-140-274-394 | Sequence 394, App |
| 45 | 44 | 50.6 | 1184 | 12 | US-10-140-471-394 | Sequence 394, App |

ALIGNMENTS

RESULT 1
US-10-236-055A-14
; Sequence 14, Application US/10236055A
; Publication No. US20030134328A1
; GENERAL INFORMATION:
; APPLICANT: Basham, Beth E.
; APPLICANT: Forsythe, Ian
; APPLICANT: Gorman, Daniel M.
; APPLICANT: Mattson, Jeanine
; APPLICANT: Moshrefi, Mehrdad
; APPLICANT: Parham, Christi
; TITLE OF INVENTION: MAMMALIAN GENES; RELATED REAGENTS
; FILE REFERENCE: DX01343K
; CURRENT APPLICATION NUMBER: US/10/236, 055A
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/317, 988
; PRIOR FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-236-055A-14
Query Match 100.0%; Score 87; DB 12; Length 329;
Best local Similarity 100.0%; Pred. No. 0.00012;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLPLLLGLGPACWA 16
Db 1 MLPLLLGLGPACWA 16
RESULT 2
US-10-085-167-2
; Sequence 2, Application US/10085167
; Publication No. US20030170781A1


```
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Lok, Si
; TITLE OF INVENTION: SECRETED PROTEIN ZACRP4
; FILE REFERENCE: 99-29
; CURRENT APPLICATION NUMBER: US/10/085,167
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: 60/141,928
; PRIOR FILING DATE: 1999-07-01
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 2
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-10-085-167-2

Query Match      100.0%; Score 87; DB 12; Length 329;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MLPLLGLGPACWA 16
      |||||
Db      1 MLPLLGLGPACWA 16

RESULT 3
US-10-236-055A-16
; Sequence 16, Application US/102336055A
; Publication No. US20030134328A1
; GENERAL INFORMATION:
; APPLICANT: Basham, Beth E.
; APPLICANT: Forsythe, Ian
; APPLICANT: Gorman, Daniel M.
; APPLICANT: Matison, Jeanine
; APPLICANT: Moshrefi, Mehrdad
; APPLICANT: Parham, Christi
; TITLE OF INVENTION: MAMMALIAN GENES; RELATED REAGENTS
; FILE REFERENCE: DX01343K
; CURRENT APPLICATION NUMBER: US/10/236,055A
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/317,988
; PRIOR FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Mus musculus
;
US-10-236-055A-16

Query Match      83.9%; Score 73; DB 12; Length 326;
Best Local Similarity 87.5%; Pred. No. 0.011;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 MLPLLGLGPACWA 16
      |||||
Db      1 MLPLLGLGPACWA 16

RESULT 4
US-09-759-130B-42
; Sequence 42, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: KIRST, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
```

```
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES.
; FILE REFERENCE: MP100-535OMNIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 42
; LENGTH: 1135
; TYPE: PRT
; ORGANISM: Mus sp.
;
US-09-759-130B-42

Query Match      56.3%; Score 49; DB 11; Length 1135;
Best Local Similarity 69.2%; Pred. No. 95;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 MLPLLGLGPAA 13
      :|||
Db      4 LLPFLGLGPGS 16

RESULT 5
US-10-120-801-40
; Sequence 40, Application US/10120801
; Publication No. US20030203843A1
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Mehraban, Fuad
; APPLICANT: Topper, James N.
; APPLICANT: Malyankar, Uriel
; APPLICANT: Wasserman, Scott
; APPLICANT: Edinger, Shlomit
; APPLICANT: Smithson, Glenda
; APPLICANT: Gunther, Erik
; APPLICANT: Komuves, Laszlo
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-340
; CURRENT APPLICATION NUMBER: US/10/120,801
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/285748
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/286068
```

```
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/286292
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/288334
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/291241
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 60/322284
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/285609
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 40
; LENGTH: 379
; TYPE: PRT
; ORGANISM: mouse
US-10-120-801-40
```

```
Query Match          55.2%; Score 48; DB 12; Length 379;
Best Local Similarity 69.2%; Pred. No. 47;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 MLPLLGLGPPAA 13
      :|||:|||:|
Db      31 LLPLWLGLMGPGA 43
```

```
RESULT 6
US-09-554-000-8
```

```
; Sequence 8, Application US/09554000
; Patent No. US20020165364A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; TITLE OF INVENTION: DETECTION OF ANALYTES
; FILE REFERENCE: 07257/042001
; CURRENT APPLICATION NUMBER: US/09/554,000
; CURRENT FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 08/818,252
; PRIOR FILING DATE: 1997-03-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 656
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-09-554-000-8
```

```
Query Match          54.0%; Score 47; DB 10; Length 656;
Best Local Similarity 83.3%; Pred. No. 1.le+02;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      2 LPPLLGLGPPAA 13
      :|||:|||:|
Db      5 VPPLLGLGAAA 16
```

```
RESULT 7
US-09-554-000-56
```

```
; Sequence 56, Application US/09554000
; Patent No. US20020165364A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; TITLE OF INVENTION: DETECTION OF ANALYTES
; FILE REFERENCE: 07257/042001
; CURRENT APPLICATION NUMBER: US/09/554,000
; CURRENT FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 08/818,252
; PRIOR FILING DATE: 1997-03-14
```

```
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-554-000-56
```

```
Query Match          51.7%; Score 45; DB 10; Length 18;
Best Local Similarity 83.3%; Pred. No. 7;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      2 LPPLLGLGPPAA 13
      :|||:|||:|
Db      5 VPPLLGLGAAA 16
```

```
RESULT 8
US-10-120-801-41
```

```
; Sequence 41, Application US/10120801
; Publication No. US20030203843A1
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Mehraban, Fuad
; APPLICANT: Topper, James N.
; APPLICANT: Malyanfar, Uriel
; APPLICANT: Wasserman, Scott
; APPLICANT: Edinger, Shlomit
; APPLICANT: Smithson, Glenda
; APPLICANT: Gunther, Erik
; APPLICANT: Komuves, Laszlo
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-340
; CURRENT APPLICATION NUMBER: US/10/120,801
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/285748
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/286068
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/286292
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/288334
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/291241
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 60/322284
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/285609
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 41
; LENGTH: 347
; TYPE: PRT
; ORGANISM: human
US-10-120-801-41
```

```
Query Match          51.7%; Score 45; DB 12; Length 347;
Best Local Similarity 69.2%; Pred. No. 1.le+02;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy      1 MLPLLGLGPPAA 13
      :|||:|||:|
Db      29 LLPLWLGLAGPGA 41
```

```
RESULT 9
US-10-120-801-4
```

```
; Sequence 4, Application US/10120801
; Publication No. US20030203843A1
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Mehraban, Fuad
; APPLICANT: Topper, James N.
; APPLICANT: Malyankar, Uriel
; APPLICANT: Wasserman, Scott
; APPLICANT: Edinger, Shlomit
; APPLICANT: Smithson, Glenda
; APPLICANT: Gunther, Erik
; APPLICANT: Komuves, Laszlo
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-340
; CURRENT APPLICATION NUMBER: US/10/120,801
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/285748
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/286068
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/286292
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/288334
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/291241
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 60/322284
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/285609
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-120-801-4

Query Match      51.7%; Score 45; DB 12; Length 378;
Best Local Similarity 69.2%; Pred. No. 1.2e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      1 MLPULLGLGPAA 13
      :|||||
Db      29 LLPMLGLAGPGA 41
```

```
RESULT 10
US-10-316-253-2
; Sequence 2, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Feng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 416
; TYPE: PRT
```

```
; ORGANISM: Rattus norvegicus
; US-10-316-253-2
```

```
Query Match      51.7%; Score 45; DB 12; Length 416;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 LPULLGLGPAA 13
      :|||||
Db      5 VPULLGLGLAA 16
```

```
RESULT 11
US-10-316-253-4
; Sequence 4, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Feng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-10-316-253-4
```

```
Query Match      51.7%; Score 45; DB 12; Length 416;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 LPULLGLGPAA 13
      :|||||
Db      5 VPULLGLGLAA 16
```

```
RESULT 12
US-10-316-253-6
; Sequence 6, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Feng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-10-316-253-6
```

```
Query Match      51.7%; Score 45; DB 12; Length 416;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

QY 2 LPLGLGPAA 13
:|||||||
Db 5 VPLGLGLGAA 16

RESULT 13

US-10-369-493-4780
; Sequence 4780, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 4780
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Burkholderia fungorum
US-10-369-493-4780

Query Match 50.6%; Score 44; DB 12; Length 242;
Best Local Similarity 64.3%; Pred. No. 1.1e+02;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 PLLGLGPACWA 16
||:|||||
Db 57 PLVVDLDPACSA 70

RESULT 14

US-10-369-493-7540
; Sequence 7540, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 7540
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
US-10-369-493-7540

Query Match 50.6%; Score 44; DB 12; Length 242;
Best Local Similarity 64.3%; Pred. No. 1.1e+02;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 PLLGLGPACWA 16
||:|||||
Db 57 PLVVDLDPACSA 70

RESULT 15
US-10-156-761-10413

; Sequence 10413, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10413
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10413

Query Match 50.6%; Score 44; DB 15; Length 394;
Best Local Similarity 55.0%; Pred. No. 1.8e+02;
Matches 11; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

QY 2 LPLGLL-----GPACW 15
||| |||||
Db 113 LPWLTGLTVIGVGPAVW 132

Search completed: January 12, 2004, 08:33:40
Job time : 4.9481 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:14 ; Search time 0.814371 Seconds
(without alignments)
831.284 Million cell updates/sec

Title: US-10-085-167-2_COPY_1_16
Perfect score: 87
Sequence: 1 MLPPLGLGPRACWA 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------------------|--------------------|
| 1 | 52 | 59.8 | 135 | 4 US-09-252-991A-21441 | Sequence 21441, A |
| 2 | 47 | 54.0 | 656 | 2 US-08-818-253-8 | Sequence 8, Appli |
| 3 | 47 | 54.0 | 656 | 3 US-08-818-252-8 | Sequence 8, Appli |
| 4 | 45 | 51.7 | 18 | 2 US-08-818-253-56 | Sequence 56, Appli |
| 5 | 45 | 51.7 | 18 | 2 US-08-818-253-61 | Sequence 61, Appli |
| 6 | 45 | 51.7 | 18 | 3 US-08-818-252-56 | Sequence 56, Appli |
| 7 | 44 | 50.6 | 318 | 1 US-08-375-962B-14 | Sequence 14, Appli |
| 8 | 44 | 50.6 | 318 | 3 US-08-940-424-3 | Sequence 3, Appli |
| 9 | 43 | 49.4 | 96 | 3 US-08-465-343A-11 | Sequence 11, Appli |
| 10 | 42 | 48.3 | 109 | 4 US-09-199-637A-389 | Sequence 389, App |
| 11 | 41 | 47.1 | 90 | 4 US-09-252-991A-19814 | Sequence 19814, A |
| 12 | 41 | 47.1 | 277 | 2 US-08-147-784-2 | Sequence 2, Appli |
| 13 | 41 | 47.1 | 277 | 3 US-08-195-967-2 | Sequence 2, Appli |
| 14 | 41 | 47.1 | 277 | 3 US-09-006-353A-12 | Sequence 12, Appli |
| 15 | 41 | 47.1 | 277 | 3 US-08-472-940-2 | Sequence 2, Appli |
| 16 | 41 | 47.1 | 277 | 4 US-09-573-986-12 | Sequence 12, Appli |
| 17 | 41 | 47.1 | 277 | 4 US-09-880-939-2 | Sequence 2, Appli |
| 18 | 41 | 47.1 | 277 | 4 US-09-804-200-2 | Sequence 2, Appli |
| 19 | 41 | 47.1 | 355 | 4 US-09-325-932A-191 | Sequence 191, App |
| 20 | 41 | 47.1 | 410 | 4 US-09-252-991A-31174 | Sequence 31174, A |
| 21 | 41 | 47.1 | 540 | 4 US-09-328-352-4514 | Sequence 4514, Ap |
| 22 | 41 | 47.1 | 1182 | 4 US-09-293-505-7 | Sequence 7, Appli |
| 23 | 41 | 47.1 | 1203 | 4 US-09-207-857-2 | Sequence 2, Appli |
| 24 | 41 | 47.1 | 1203 | 4 US-09-293-505-2 | Sequence 2, Appli |
| 25 | 41 | 47.1 | 1311 | 2 US-08-540-406-4 | Sequence 4, Appli |
| 26 | 41 | 47.1 | 1311 | 3 US-08-656-055-4 | Sequence 4, Appli |
| 27 | 41 | 47.1 | 1311 | 3 US-08-954-668-4 | Sequence 4, Appli |

| | | | | | |
|----|------|------|------|------------------------|-------------------|
| 28 | 41 | 47.1 | 1311 | 4 US-08-918-658-4 | Sequence 4, Appli |
| 29 | 41 | 47.1 | 1311 | 4 US-09-724-631-4 | Sequence 4, Appli |
| 30 | 41 | 47.1 | 1311 | 5 PCT-US95-13233-4 | Sequence 4, Appli |
| 31 | 40.5 | 46.6 | 176 | 4 US-09-252-991A-26925 | Sequence 26925, A |
| 32 | 40 | 46.0 | 218 | 4 US-09-252-991A-25415 | Sequence 25415, A |
| 33 | 40 | 46.0 | 584 | 2 US-08-845-295A-2 | Sequence 2, Appli |
| 34 | 40 | 46.0 | 584 | 3 US-09-140-933-2 | Sequence 2, Appli |
| 35 | 40 | 46.0 | 584 | 3 US-09-146-661-2 | Sequence 2, Appli |
| 36 | 40 | 46.0 | 584 | 3 US-09-150-515-2 | Sequence 2, Appli |
| 37 | 40 | 46.0 | 986 | 3 US-08-872-757-4 | Sequence 4, Appli |
| 38 | 40 | 46.0 | 986 | 4 US-09-850-048A-4 | Sequence 4, Appli |
| 39 | 40 | 46.0 | 999 | 1 US-08-252-626A-2 | Sequence 2, Appli |
| 40 | 39.5 | 45.4 | 495 | 4 US-09-252-991A-19330 | Sequence 19330, A |
| 41 | 39 | 44.8 | 182 | 4 US-09-134-001C-5022 | Sequence 5022, Ap |
| 42 | 39 | 44.8 | 192 | 3 US-08-486-099-107 | Sequence 107, App |
| 43 | 39 | 44.8 | 192 | 3 US-08-360-107A-117 | Sequence 117, App |
| 44 | 39 | 44.8 | 192 | 3 US-08-484-223B-107 | Sequence 107, App |
| 45 | 39 | 44.8 | 192 | 3 US-08-919-597-107 | Sequence 107, App |

ALIGNMENTS

```
RESULT 1
US-09-252-991A-21441
; Sequence 21441, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252, 991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074, 788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094, 190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21441
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-21441

Query Match      59.8%; Score 52; DB 4; Length 135;
Best Local Similarity 56.2%; Pred. No. 0.58;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      1 MLPPLGLGPRACWA 16
DB      23 VLPMLALAVGLCWA 38

RESULT 2
US-08-818-253-8
; Sequence 8, Application US/08818253
; Patent No. 5998204
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
```

```
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/818,253
FILING DATE: 14-MAR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07257/043001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 656 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-818-253-8
```

```
Query Match          54.0%; Score 47; DB 2; Length 656;
Best Local Similarity 83.3%; Pred. No. 19;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      2 LPULLGLGPAA 13
       :|||||||
Db      5 VPULLGLLGAAA 16
```

```
RESULT 3
US-08-818-252-8
Sequence 8, Application US/08818252B
Patent No. 6197928
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
APPLICANT: Miyawaki, Atsushi
TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
TITLE OF INVENTION: DETECTION OF ANALYTES
FILE REFERENCE: 07257/042001
CURRENT APPLICATION NUMBER: US/08/818,252B
CURRENT FILING DATE: 1997-03-14
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 656
TYPE: PRT
ORGANISM: Aequorea victoria
US-08-818-252-8
```

```
Query Match          54.0%; Score 47; DB 3; Length 656;
Best Local Similarity 83.3%; Pred. No. 19;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      2 LPULLGLGPAA 13
       :|||||||
Db      5 VPULLGLLGAAA 16
```

```
RESULT 4
US-08-818-253-56
Sequence 56, Application US/08818253
Patent No. 5998204
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
APPLICANT: Miyawaki, Atsushi
TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
TITLE OF INVENTION: DETECTION OF ANALYTES
NUMBER OF SEQUENCES: 61
```

```
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/818,253
FILING DATE: 14-MAR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07257/043001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-818-253-56
```

```
Query Match          51.7%; Score 45; DB 2; Length 18;
Best Local Similarity 83.3%; Pred. No. 0.8;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      2 LPULLGLGPAA 13
       :|||||||
Db      5 VPULLGLLGAAA 16
```

```
RESULT 5
US-08-818-253-61
Sequence 61, Application US/08818253
Patent No. 5998204
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
APPLICANT: Miyawaki, Atsushi
TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
TITLE OF INVENTION: DETECTION OF ANALYTES
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/818,253
FILING DATE: 14-MAR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
```

```

; REFERENCE/DOCKET NUMBER: 07257/043001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-818-253-61

Query Match          51.7%; Score 45; DB 2; Length 18;
Best local Similarity 83.3%; Pred. No. 0.8;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 LPLLLGLGPAA 13
       :|||||
Db      5 VPLLLGLGLAA 16

RESULT 6
US-08-818-252-56
; Sequence 56, Application US/08818252B
; Patent No. 6197928
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; TITLE OF INVENTION: DETECTION OF ANALYTES
; FILE REFERENCE: 07257/042001
; CURRENT APPLICATION NUMBER: US/08/818,252B
; CURRENT FILING DATE: 1997-03-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-08-818-252-56

Query Match          51.7%; Score 45; DB 3; Length 18;
Best local Similarity 83.3%; Pred. No. 0.8;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 LPLLLGLGPAA 13
       :|||||
Db      5 VPLLLGLGLAA 16

RESULT 7
US-08-375-962B-14
; Sequence 14, Application US/08375962B
; Patent No. 5731195
; GENERAL INFORMATION:
; APPLICANT: SIMON, ANDRAS; HELLMAN, ULF; WERNSTEDT,
; APPLICANT: CHRISTER, ERIKSSON, ULF.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecule
; TITLE OF INVENTION: Which Codes for A 32 kDa Protein Having 11-CIS Retinol
; TITLE OF INVENTION: Dehydrogenase Activity, and Which Associates With P63,
; TITLE OF INVENTION: a Portion of a Retinol Binding Protein Receptor
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect (ASCII standard)
```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/375,962B
; FILING DATE: 20-January-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/258,418
; FILING DATE: 6-October-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Pasqualini, Patricia A.
; REGISTRATION NUMBER: 34,894
; REFERENCE/DOCKET NUMBER: LUD 5372.1 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 318 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; FEATURE:
; NAME/KEY: p32,11-cis retinol dehydrogenase
US-08-375-962B-14

Query Match          50.6%; Score 44; DB 1; Length 318;
Best local Similarity 71.4%; Pred. No. 25;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2 LPLLLGLGPACW 15
       :|||||
Db      3 LPLLLGVLLMAALW 16

RESULT 8
US-08-940-424-3
; Sequence 3, Application US/08940424A
; Patent No. 6171837
; GENERAL INFORMATION:
; APPLICANT: Blaner, William S.
; APPLICANT: Zott, Roseann P.
; APPLICANT: Gamble, Mary V.
; APPLICANT: Mertz, James R.
; TITLE OF INVENTION: POTENT INHIBITORS OF HUMAN 9-cis RETINOL DEHYDROGENASE
; FILE REFERENCE: 0575/54544
; CURRENT APPLICATION NUMBER: US/08/940,424A
; CURRENT FILING DATE: 1997-09-29
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 318
; TYPE: PRT
; ORGANISM: rat
US-08-940-424-3

Query Match          50.6%; Score 44; DB 3; Length 318;
Best local Similarity 71.4%; Pred. No. 25;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2 LPLLLGLGPACW 15
       :|||||
Db      3 LPLLLGVLLMAALW 16

RESULT 9
US-08-465-343A-11
; Sequence 11, Application US/08465343A
; Patent No. 6214968
; GENERAL INFORMATION:
; APPLICANT: Wyler, David J.
; APPLICANT: Prakash, Sadhana
; APPLICANT: Zhang, Xiaoping
; TITLE OF INVENTION: FSF-1 AND THE EARLY DETECTION
; TITLE OF INVENTION: OF FIBROSIS
```



```

; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,343A
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/395,674
; FILING DATE: 28-FEB-1995
; APPLICATION NUMBER: 08/152,904
; FILING DATE: 15-NOV-1993
; APPLICATION NUMBER: 07/840,426
; FILING DATE: 24-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Weikiejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 00398/096002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-465-343A-11

Query Match 49.4%; Score 43; DB 3; Length 96;
Best Local Similarity 90.0%; Pred. No. 9.9;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLPLLGLLG 10
Db 1 MLPLLGLLYG 10

RESULT 10
US-09-199-637A-389
; Sequence 389, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Teongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 389
; LENGTH: 109
```

```

; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-199-637A-389

Query Match 48.3%; Score 42; DB 4; Length 109;
Best Local Similarity 60.0%; Pred. No. 16;
Matches 9; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

QY 7 GLGPA-----CW 15
Db 71 GLGPAETPLSICW 85

RESULT 11
US-09-252-991A-19814
; Sequence 19814, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19814
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-19814

Query Match 47.1%; Score 41; DB 4; Length 90;
Best Local Similarity 69.2%; Pred. No. 19;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LPLLGLGPAC 14
Db 30 LPLLMFLGLAAC 42

RESULT 12
US-08-147-784-2
; Sequence 2, Application US/08147784
; Patent No. 5821332
; GENERAL INFORMATION:
; APPLICANT: Godfrey, Wayne
; APPLICANT: Buck, David
; APPLICANT: Engleman, Edgar G.
; TITLE OF INVENTION: RECEPTOR ON THE SURFACE OF ACTIVATED
; TITLE OF INVENTION: CD4+ T-CELLS: ACT-4
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Knourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/147,784
; FILING DATE: 03-NOV-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
```

REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 05490A-220
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 277 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-147-784-2

Query Match 47.1%; Score 41; DB 2; Length 277;
Best Local Similarity 75.0%; Pred. No. 63;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LPLLLGLGPAA 13
| : ||||| |
DB 221 LGLVLGLGPLA 232

RESULT 13

US-08-195-967-2
Sequence 2, Application US/08195967
Patent No. 6242566
GENERAL INFORMATION:
APPLICANT: Godfrey, Wayne
APPLICANT: Engleman, Edgar G.
TITLE OF INVENTION: LIGAND (ACT-4-L) TO A RECEPTOR ON THE SURFACE OF ACTIVATED
TITLE OF INVENTION: CD4+ T-CELLS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/195,967
FILING DATE: 10-FEB-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 05490A-230
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 277 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-195-967-2

Query Match 47.1%; Score 41; DB 3; Length 277;
Best Local Similarity 75.0%; Pred. No. 63;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LPLLLGLGPAA 13
| : ||||| |
DB 221 LGLVLGLGPLA 232

RESULT 14

US-09-006-353A-12
Sequence 12, Application US/09006353A
Patent No. 6261801
GENERAL INFORMATION:
APPLICANT: WEI, YING-FEI
APPLICANT: YU, GUO-LIANG
APPLICANT: GENTZ, REINER
APPLICANT: RUBEN, STEVEN
TITLE OF INVENTION: TUMOR NECROSIS FACTOR RECEPTOR 5
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/006,353A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF341
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 277 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-006-353A-12

Query Match 47.1%; Score 41; DB 3; Length 277;
Best Local Similarity 75.0%; Pred. No. 63;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LPLLLGLGPAA 13
| : ||||| |
DB 221 LGLVLGLGPLA 232

RESULT 15

US-08-472-940-2
Sequence 2, Application US/08472940
Patent No. 6277962
GENERAL INFORMATION:
APPLICANT: Godfrey, Wayne
APPLICANT: Buck, David
APPLICANT: Engleman, Edgar G.
TITLE OF INVENTION: RECEPTOR ON THE SURFACE OF ACTIVATED
TITLE OF INVENTION: CD4+ T-CELLS: ACT-4
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: Patentin Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,940
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/147,784
; FILING DATE: 03-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Willaim M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 05490A-220
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 277 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-472-940-2
    
```

```

Query Match          47.1%; Score 41; DB 3; Length 277;
Best Local Similarity 75.0%; Pred. No. 63;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 LPVLGLGPPA 13
      ||:|||||||
Db      221 LGLVLGLGPPA 232
    
```

Search completed: January 12, 2004, 08:20:07
 Job time : 2.81437 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:09 ; Search time 8.27744 Seconds
(without alignments)
1661.397 Million cell updates/sec

Title: US-10-085-167-2_COPY_17_159
Perfect score: 745
Sequence: 1 LGPTPGGSSELSAFAAAR.....PHYALGAPGATFSGYLVYAD 143

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 191 | 25.6 | 244 | 2 JC4708 | gelatin-binding 28 |
| 2 | 189 | 25.4 | 219 | 2 T14782 | hypothetical prote |
| 3 | 168.5 | 22.6 | 253 | 1 C1HUQB | complement subcomp |
| 4 | 163.5 | 21.9 | 253 | 1 S49158 | complement subcomp |
| 5 | 160 | 21.5 | 245 | 1 C1HUQC | complement subcomp |
| 6 | 156 | 20.9 | 245 | 2 S19018 | complement subcomp |
| 7 | 152.5 | 20.5 | 246 | 2 S29328 | complement subcomp |
| 8 | 150 | 20.1 | 224 | 2 A60032 | cerebellin-like gl |
| 9 | 149.5 | 20.1 | 253 | 2 I49560 | complement C1q B c |
| 10 | 149 | 20.0 | 193 | 2 A37873 | cerebellin precurs |
| 11 | 148.5 | 19.9 | 674 | 2 S23297 | collagen alpha 1(X |
| 12 | 148 | 19.9 | 674 | 2 S13301 | collagen alpha 1(X |
| 13 | 147 | 19.7 | 245 | 1 C1HUQA | complement subcomp |
| 14 | 145 | 19.5 | 680 | 1 CGHU1D | collagen alpha 1(X |
| 15 | 144.5 | 19.4 | 215 | 2 C48150 | hibernation-relate |
| 16 | 144 | 19.3 | 1228 | 2 A57384 | multimerin, endoth |
| 17 | 139.5 | 18.7 | 635 | 2 A57131 | collagen alpha 2(V |
| 18 | 139.5 | 18.7 | 744 | 1 A34246 | collagen alpha 1(V |
| 19 | 139 | 18.7 | 423 | 2 A55797 | collagen precursor |
| 20 | 137.5 | 18.5 | 744 | 2 S15435 | collagen alpha 1(V |
| 21 | 136.5 | 18.3 | 215 | 2 B48150 | hibernation-relate |
| 22 | 136 | 18.3 | 743 | 1 S23779 | collagen alpha 1(V |
| 23 | 135.5 | 18.2 | 170 | 2 B57131 | collagen alpha 2(V |
| 24 | 134.5 | 18.1 | 744 | 1 S23298 | collagen alpha 1(V |
| 25 | 132 | 17.7 | 680 | 2 S31216 | collagen alpha 1(X |
| 26 | 121.5 | 16.3 | 992 | 2 T08772 | hypothetical prote |
| 27 | 113.5 | 15.2 | 196 | 2 A48150 | hibernation-relate |
| 28 | 76 | 10.2 | 366 | 2 T52655 | thiosulfate sulfur |
| 29 | 76 | 10.2 | 483 | 2 E75584 | conserved hypothet |

| | | | | | |
|----|------|------|-----|----------|--------------------|
| 30 | 75 | 10.1 | 272 | 2 B38453 | anaerobic sulfite |
| 31 | 74 | 9.9 | 272 | 2 AD0825 | anaerobic sulfite |
| 32 | 74 | 9.9 | 564 | 2 E87631 | acyl-CoA dehydroge |
| 33 | 73 | 9.8 | 413 | 2 S22578 | translation initia |
| 34 | 73 | 9.8 | 789 | 2 S33056 | probable infected |
| 35 | 71 | 9.5 | 748 | 2 T32654 | hypothetical prote |
| 36 | 70 | 9.4 | 314 | 2 T35241 | hypothetical prote |
| 37 | 69 | 9.3 | 303 | 2 B83336 | hypothetical prote |
| 38 | 69 | 9.3 | 355 | 2 S52022 | translation initia |
| 39 | 69 | 9.3 | 413 | 2 S55898 | translation initia |
| 40 | 69 | 9.3 | 413 | 2 S52018 | translation initia |
| 41 | 69 | 9.3 | 413 | 2 S52019 | translation initia |
| 42 | 69 | 9.3 | 413 | 2 S52023 | translation initia |
| 43 | 69 | 9.3 | 571 | 2 H83696 | DNA mismatch repai |
| 44 | 69 | 9.3 | 797 | 2 A87247 | probable secreted |
| 45 | 68.5 | 9.2 | 333 | 2 B84476 | probable Athila re |

ALIGNMENTS

RESULT 1

JC4708
gelatin-binding 28K protein precursor - human
N/Alternate names: adipose specific collagen-like factor
C/Species: Homo sapiens (man)
C/Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 20-Sep-1999
C/Accession: JC4708; JC4944
R/Maeda, K.; Okubo, K.; Shimomura, I.; Funahashi, T.; Matsuzawa, Y.; Matsubara, K.
Biochem. Biophys. Res. Commun. 221, 286-289, 1996
A/Title: cDNA cloning and expression of a novel adipose specific collagen-like factor,
A/Reference number: JC4708; MUID:96224171; PMID:8619847
A/Accession: JC4708
A/Molecule type: mRNA
A/Residues: 1-244 <MAE>
A/Cross-references: DDBJ:D45371; NID:G871886; PIDN:BAA08227.1; PID:G871887
A/Experimental source: adipose tissue
R/Nakano, Y.; Tobe, T.; Choi-Miura, N.H.; Mazda, T.; Tomita, M.
J. Biochem. 120, 803-812, 1996
A/Title: Isolation and characterization of GBP28, a novel gelatin-binding protein purifi
A/Reference number: JC4944; MUID:97103474; PMID:8947845
A/Accession: JC4944
A/Molecule type: protein
A/Residues: 19-38;93-100;101-112;135-149;173-178 <NAK>
C/Comment: This protein is an endogenous factor that binds with a collagen-like domain.
C/Genetics:
A/Gene: apM1
C/Superfamily: unassigned collagens; complement C1q carboxyl-terminal homology
C/Keywords: adipose tissue; glycoprotein; hydroxyproline
F/1-18/Domain: signal sequence #status predicted <SIG>
F/19-244/Product: gelatin-binding 28kDa protein #status experimental <MAT>
F/42-107/Region: collagen-like
F/114-241/Domain: complement C1q carboxyl-terminal homology <C1Q>
F/95/Modified site: 4-hydroxyproline (Pro) #status experimental
F/230/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 25.6%; Score 191; DB 2; Length 244;
Best local Similarity 31.8%; Pred. No. 6.3e-12;
Matches 47; Conservative 29; Mismatches 54; Indels 18; Gaps 6;

| | | | |
|----|-----|--|-----|
| QY | 5 | PGPSSSELSAFAAARTTPLEGTSSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAYFFSF | 64 |
| DB | 104 | PGEQAYVYRSASFVGLETVYT-IPNWPVIRFTKIFYNQONHYDSTGKFCNIPGLYFYAY | 162 |
| QY | 65 | TAGKAPH-----KSLSVMLVRNRDEVQALAFDEQRRPGARRAASQASMLQLDYGDVWLR | 119 |
| DB | 163 | -----HITVYMKDVKVSIFK-KDKAMLFTYDQYQENNVDQ-ASGSVLLHLEVGDQVWLQ | 214 |
| QY | 120 | LHGAPH---YALGAPGATFSGYLVYAD | 143 |
| DB | 215 | VYGEGERNGLYADNDNDSTFTGFLLYHD | 242 |

RESULT 2

TI14782
hypothetical protein DKFZp586B0621.1 - human (fragment)
A;Accession: B90315
A;Molecule type: protein
C;Species: Homo sapiens (man)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Sep-2000
C;Accession: TI14782
R;Ottewaeldeir, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, August 1999
A;Reference number: Z18184
A;Accession: TI14782
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-219 <OTT>
A;Cross-references: EMBL:AL110261
A;Experimental source: adult uterus; clone DKFZp586B0621
C;Genetics:
A;Note: DKFZp586B0621.1
C;Superfamily: complement C1q carboxyl-terminal homology

Query Match 25.4%; Score 189; DB 2; Length 219;
Best Local Similarity 35.3%; Pred. No. 8.9e-12;
Matches 53; Conservative 28; Mismatches 57; Indels 12; Gaps 6;

QY 2 GPTPGGSSSEL--RSAFSAART-TPLEGTSEMAVTFDKYVYVNIIGDFDVATGQRCRVPG 58
Db 66 GPTGPAGECSVPSPRSFAKSSESRRVPPSPDAPLPFDRLVNEQGHYDAVTGKFTCOVPG 125
QY 59 AYFFSFTAGKAPHKSLSVMLVNRDEVOAL--AFDEQRPRGARRAASQSAAMLQLDYGDV 116
Db 126 VYFPAVHA-TVYRASLQFDLVKNGESIASFFQFFGWPKPRA--SLSGAMVRLBEDQV 181
QY 117 WLRLHGAPH---YALGAPGATFSGYLVYAD 143
Db 182 WVQVGVDYIGIYASIKTDSFGFLVYSD 211

RESULT 3

C1HUQB

complement subcomponent C1q chain B precursor [validated] - human
N;Alternate names: complement subcomponent C1q beta chain

C;Species: Homo sapiens (man)
C;Date: 22-May-1981 #sequence_revision 31-May-1996 #text_change 08-Dec-2000
C;Accession: B23422; A23422; B90304; A90301; B90315; A03206
R;Reid, K.B.M.

Biochem. J. 231, 729-735, 1985

A;Title: Molecular cloning and characterization of the complementary DNA and gene coding
A;Reference number: A23422; MUID:86076906; PMID:3000358

A;Accession: B23422

A;Molecule type: DNA

A;Residues: 'HS', 1-32 <REI>

A;Note: the authors translated the codon AGT for the second position as Arg; they were u

A;Accession: A23422

A;Molecule type: mRNA

A;Residues: 28-253 <RE1>

A;Cross-references: EMBL:X03084

A;Note: the authors translated the codon ACA for residue 46 as Ile

R;Reid, K.B.M.

Biochem. J. 179, 367-371, 1979

A;Title: Complete amino acid sequences of the three collagen-like regions present in sub

A;Reference number: A90304; MUID:80020137; PMID:486087

A;Accession: B90304

A;Molecule type: protein

A;Residues: 'E', 29-84, 'D', 86-99, 'P', 101-135 <RE5>

R;Reid, K.B.M.; Thompson, B.O.P.

Biochem. J. 173, 863-868, 1978

A;Title: Amino acid sequence of the N-terminal 108 amino acid residues of the B chain of

A;Reference number: A90301; MUID:79041552; PMID:708376

A;Accession: A90301

A;Molecule type: protein

A;Residues: 28-99, 'P', 101-195 <RE3>

R;Reid, K.B.M.; Gagnon, J.; Frampton, J.

Biochem. J. 203, 559-569, 1982

A;Title: Completion of the amino acid sequences of the A and B chains of subcomponent C1

A;Reference number: A90315; MUID:82283890; PMID:6981411

A;Accession: B90315

A;Molecule type: protein

A;Residues: 136-253 <RE4>

A;Note: 176-Glx may also be present

C;Comment: The first component of complement is a calcium-dependent complex of the three

vation of C1r (enzyme), C1s (proenzyme), and the other eight components of complement.

C;Comment: The C1q subcomponent is composed of nine subunits, six of which are disulfid

(see PIR:C1HUQC) chain. Equimolar amounts of the A, B, and C chains are found after red

C;Genetics:

A;Gene: GDB:C1QB

A;Cross-references: GDB:119043; OMIM:120570

A;Map position: 1p36.3-1p34.1

C;Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal ho

C;Keywords: complement pathway; glycoprotein; heterodimer; hydroxyllysine; hydroxyprolin

F;1-27/Domain: signal sequence #status predicted <SIG>

F;28-253/Product: complement subcomponent C1q chain B #status experimental <MAT>

F;33-116/Domain: collagenous, triple helix <COL>

F;123-249/Domain: complement C1q carboxyl-terminal homology <C1Q>

F;28/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experime

F;31/Disulfide bonds: interchain (to chain A-26) #status experimental

F;35,38,41,53,56,65,83,86,101,104,107/Modified site: 4-hydroxyproline (Pro) #status exp

F;59,62,77,92,98,110/Modified site: 5-hydroxylysine (Lys) #status experimental

F;59,62,98,110/Binding site: carbohydrate (Lys) (covalent) #status experimental

Query Match 22.6%; Score 168.5; DB 1; Length 253;
Best Local Similarity 27.7%; Pred. No. 1.3e-09;
Matches 44; Conservative 34; Mismatches 62; Indels 19; Gaps 4;

QY 1 LGPTPGGS-----SELRSASARTTPLEGTSEMAVTFDKYVYVNIIGDFD 46
Db 95 MGPKGGPAGPAGPBGKESGDYKATOKIAFSATRTINVPRLRRDQITREDHVITNMNNYE 154
QY 47 VATGQRCRVPGAYFFSFTAGKAPHKSLSVMLVNRDEVOALAFDEQRPRGARRAASQSA 106
Db 155 PRSGKFTCKVPGLYYFTYHA--SSRGNLCVNLMRGERAOKVTFCDYAVNTFQVTTGGM 212
QY 107 MLQLDYGDTVWLRLHGAPHYALGAPGAT--FSGYLVYAD 143
Db 213 VLKLEGEENVFLQATD-KNSLLGMEGANSIFSGFLLPD 250

RESULT 4

S49158

complement protein C1q beta chain precursor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 20-Aug-1999

C;Accession: S49158

R;Schwaeble, W.; Petry, F.; Loos, M.

submitted to the EMBL Data Library, March 1993

A;Description: cDNA cloning and expression of the mRNA encoding the B-chain of rat C1q.

A;Reference number: S49158

A;Accession: S49158

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-253 <SCH>

C;Cross-references: EMBL:X71127; NID:G510191; PIDN:CA50440.1; PID:G510192

C;Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal ho

F;121-249/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 21.9%; Score 163.5; DB 2; Length 253;
Best Local Similarity 29.2%; Pred. No. 4.2e-09;
Matches 45; Conservative 36; Mismatches 58; Indels 15; Gaps 7;

QY 2 GPTPGP-----GSSELSRS----AFSARTTPLEGTSEMAVTFDKYVYVNIIGDFDVATGQF 52
Db 100 GP-PGPRPGKGGSDYKATQKVAFSALRTVNSALRPNOAIRFEKVTITVNDNYEPRSGKF 158
QY 53 RCRVPGAYFFSFTAGKAPHKSLSVMLVNRDEVQ-ALAFDEQRPRGARRAASQSAAMLQD 111
Db 159 TCKVPGLYYFTYHASSRGNLCVNLVIRGRDRDMQKVLTFCDYAO-NTFQVTTGGVVLKLE 217
QY 112 YGDTVWLRLHGAPHYALGAPGAT--FSGYLVYAD 143

Db 218 QEEVHLQATD-KNSLLGVEGANSIFTGFLFPD 250

RESULT 5
C1HUQC

complement subcomponent C1q chain C precursor - human

N/Alternate names: complement subcomponent C1q gamma chain

C/Species: Homo sapiens (man)

C/Date: 22-May-1981 #sequence_revision 31-May-1996 #text_change 22-May-1998

C/Accession: S14351; A03207

R/Sellar, G.C.; Blake, D.J.; Reid, K.B.M.

Biochem. J. 274, 481-490, 1991

A/Title: Characterization and organization of the genes encoding the A-, B- and C-chains

A/Reference number: S14350; MUID:91174759; PMID:1706597

A/Accession: S14351

A/Status: not compared with conceptual translation

A/Molecule type: DNA

A/Residues: 1-245 <SEL>

R/Reid, K.B.M.

Biochem. J. 179, 367-371, 1979

A/Title: Complete amino acid sequences of the three collagen-like regions present in sub

A/Reference number: A90304; MUID:80020137; PMID:486087

A/Accession: A03207

A/Molecule type: protein

A/Residues: 29-56,'P',58-65,'K',67-71,'P',73-83,'K',85-86,'D',88-89,'N',91-122 <REI>

C/Comment: The first component of complement is a calcium-dependent complex of the three

ivation of C1r (enzyme), C1s (proenzyme), and the other eight components of complement.

C/Comment: The C1q subcomponent is composed of nine subunits, six of which are disulfide

dimers of the C chain. Equimolar amounts of the A, B, and C chains are found after redu

C/Genetics:

A/Gene: GDB:C1QG

A/Cross-references: GDB:128132; OMIM:120575

A/Map position: 1p36.3-1p34.1

A/Introns: 60/3

C/Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom

C/Keywords: complement pathway; glycoprotein; homodimer; hydroxyllysine; hydroxyproline;

F;1-28/Domain: signal sequence #status predicted <SIG>

F;29-245/Product: complement subcomponent C1q chain B #status predicted <MAT>

F;31-114/Domain: collagenous, triple helix <COL>

F;121-244/Domain: complement C1q carboxyl-terminal homology <CIQ>

F;32/Disulfide bonds: interchain #status experimental

F;36,39,42,45,54,63,81,93,96,99,105/Modified site: 4-hydroxyproline (Pro) #status experi

F;57,72,75/Modified site: 5-hydroxyllysine (Lys) #status experimental

F;75/Binding site: carbohydrate (Lys) (covalent) #status experimental

Query Match

21.5%; Score 160; DB 1; Length 245;

Best Local Similarity 28.8%; Pred. No. 9.1e-09;

Matches 44; Conservative 28; Mismatches 63; Indels 18; Gaps 6;

QY 2 GPTPGPG-----SSELSAFAAARTTPLEGTSEMAVTFDKVYVNIIGDPAVATGQFR 53

Db 100 GPMGIPGEPGEGRYKQKQSVFTVTRQTHQPPAPNSLIRFNNAVLTNPQGDYDTSTGKFT 159

QY 54 CRVPGAYFFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLOLDY 113

Db 160 CKVPGLYFYVYHASHT--ANLCVLLYRSQVAVTFPGHSTK--TNQVNSGGVLLRLQVG 214

QY 114 DTVWLRLHGAPHY-ALGAPG--ATFSGYLVYAD 143

Db 215 EEVWLAVN--DYIDMVGIGSDSVFSGFLFPD 245

RESULT 6

S19018

complement subcomponent C1q chain A precursor - mouse

C/Species: Mus musculus (house mouse)

C/Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Aug-1999

C/Accession: S19018

R/Petry, F.; Reid, K.B.M.; Loos, M.

J. Immunol. 147, 3988-3993, 1991

A/Title: Gene expression of the A- and B-chain of mouse C1q in different tissues and the

A/Reference number: S19018; MUID:92043789; PMID:1940381

A/Accession: S19018

A/Molecule type: mRNA

A/Residues: 1-245 <PET>

A/Cross-references: EMBL:X58861; NID:g50226; PIDN:CAA41664.1; PID:g50227

C/Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal ho

F;116-243/Domain: complement C1q carboxyl-terminal homology <CIQ>

Query Match

20.9%; Score 156; DB 2; Length 245;

Best Local Similarity 31.4%; Pred. No. 2.3e-08;

Matches 48; Conservative 15; Mismatches 58; Indels 32; Gaps 4;

QY 5 PGPSSSELSAFAAARTTPLEGTSEMAVTFDKVYVNIIGDPAVATGQFRVPGAYFFSF 64

Db 106 PGNIRDQPRPAFSAIRONPM--TLGNVVFDPKVLTNQESPYQNTGRFICAVGFFYYNF 163

QY 65 -----TAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLOL 110

Db 164 QVISKWDLCLFIKSSGGQPRDSLFSNTNNKGLFQVLA-----GGTVLQL 209

QY 111 DYGDVWLRLHGAPH--YALGAPGATFSGYLVY 141

Db 210 RRGDEWIEKDPKGRITYGTEADSIFFSGFLIF 242

RESULT 7

S29328

complement subcomponent C1q chain C - mouse

C/Species: Mus musculus (house mouse)

C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Aug-1999

C/Accession: S29328

R/Petry, F.; Reid, K.B.M.; Loos, M.

Eur. J. Biochem. 209, 129-134, 1992

A/Title: Isolation, sequence analysis and characterization of cDNA clones coding for ch

ecerebellin.

A/Reference number: S29328; MUID:93011118; PMID:1396691

A/Accession: S29328

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-246 <PET>

A/Cross-references: EMBL:X66295; NID:g50228; PIDN:CAA46993.1; PID:g50229

C/Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal ho

F;122-245/Domain: complement C1q carboxyl-terminal homology <CIQ>

Query Match

20.5%; Score 152.5; DB 2; Length 246;

Best Local Similarity 28.8%; Pred. No. 5.3e-08;

Matches 44; Conservative 29; Mismatches 59; Indels 21; Gaps 5;

QY 2 GPTPGPG-----SSELSAFAAARTTPLEGTSEMAVTFDKVYVNIIGDPAVATGQFR 56

Db 104 GPPGEPGVEGRYKQKQSVFTVTRQTHQPPAPNSLIRFNNAVLTNPQGDYDTSTGKFT 163

QY 57 PGAYFF---SFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLOLDY 112

Db 164 PGLYFFVYVYTSHTANLVCVHLNLTNLARVASFCD-----HMFNSKQVSSGALLRLQR 214

QY 113 GDTVWLRLHGAPHYALGAPG--ATFSGYLVYAD 143

Db 215 GDEVWLSVNDY-NGMWGIEGSNSVFSGFLFPD 246

RESULT 8

A60032

cerebellin-like glycoprotein - rat

C/Species: Rattus norvegicus (Norway rat)

C/Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 31-Mar-2000

C/Accession: A60032

R/Wada, C.; Ohtani, H.

Brain Res. Mol. Brain Res. 9, 71-77, 1991

A/Title: Molecular cloning of rat cerebellin-like protein cDNA which encodes a novel me

A/Reference number: A60032; MUID:91203483; PMID:1850079

A/Accession: A60032

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A;Accession: S77711
A;Molecule type: protein
A;Residues: 104-112,'X',114-117;453-466 <NIN2>
C;Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
C;Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer; hydroxyproline;
F;1-18/Domain: signal sequence #status predicted <SIG>
F;547-673/Domain: complement C1q carboxyl-terminal homology <C1Q>
F;453,456/Modified site: hydroxyproline (Pro) #status experimental
F;611/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.9%; Score 148.5; DB 2; Length 674;
Best Local Similarity 27.6%; Pred. No. 4.3e-07;
Matches 47; Conservative 23; Mismatches 65; Indels 35; Gaps 5;

QY 2 GPTPGGSSELRSFAARTTPLEGTSEM-----AV 32
Db 504 GPPGPPGQSTIPEGYKGESRELSGNSFMKAGANQALTGMPVSAFTVILSKAYPGATVP 563

QY 33 TFDKYYVNIIGDFDVATGQFRCRVPGAYFFSFTAGKAPHKSLSVMLVRNRDEVQALAFDE 92
Db 564 KFDKILYNRQOHYDPTGTGIFTCRIPLGLYFSYHV-HAKGTNVWVALYKNGSPVM-YTYDE 621

QY 93 QRRPGARRAASQASAMLQLDYGDYVWLRHLGAPH---YALGAPGATFSGYL 139
Db 622 YQK-GYLDQASGSAVIDLMENDQVWLQLPNSESNGLYSSEYVHSSFSGFL 670

RESULT 12
S13301
collagen alpha 1(X) chain precursor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 21-Nov-1993 #sequence_revision 23-Feb-1996 #text_change 13-Aug-1999
C;Accession: S13301
R;Thomas, J.T.; Kwan, A.P.L.; Grant, M.E.; Boot-Handford, R.P.
Biochem. J. 273, 141-148, 1991
A;Title: Isolation of cDNAs encoding the complete sequence of bovine type X collagen. Ev
A;Reference number: S13301; MUID:91113131; PMID:1703407
A;Accession: S13301
A;Molecule type: mRNA
A;Residues: 1-674 <THO>
A;Cross-references: EMBL:X53556; NID:g263; PIDN:CAA37624.1; PID:g264
A;Genetics:
A;Gene: COL10A1
C;Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
C;Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-674/Product: collagen alpha 1(X) chain #status predicted <MAT>
F;547-673/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 19.9%; Score 148; DB 2; Length 674;
Best Local Similarity 31.5%; Pred. No. 4.9e-07;
Matches 52; Conservative 19; Mismatches 64; Indels 30; Gaps 7;

QY 2 GPTPGGSSELRSFAARTTPL---EGTSEMAVT-----FDKVVV 39
Db 511 GPPGPPGQVALPEDFVKAGORPFSVANOGVTGMPVSAFTVILSKAYPAIGTPIPDKILY 570

QY 40 NIGGDFDVATGQFRCRVPGAYFFSFTAG-KAPHKSLSVMLVRNRDEVQALAFDEQRRPGA 98
Db 571 NKQOHYDPTGTGIFTCRIPIGLYFSYHIVKGTG--AMVGLYKNGTPVM-YTYDEYIK-GY 626

QY 99 RRAASQASAMLQLDYGDYVWLRHLGAPHYALGAP---GATFSGYLV 140
Db 627 LDQASGSAVIDLTENDQVWLQLPNAGSNGLYSSEYVHSSFSGFLV 671

RESULT 13
C1HUQA
complement subcomponent C1q chain A precursor [validated] - human
N;Alternate names: complement subcomponent C1q alpha chain
C;Species: Homo sapiens (man)
C;Date: 24-Apr-1984 #sequence_revision 31-May-1996 #text_change 08-Dec-2000
C;Accession: S14350; A90304; A90315; A03205

R;Sellar, G.C.; Blake, D.J.; Reid, K.B.M.
Biochem. J. 274, 481-490, 1991
A;Title: Characterization and organization of the genes encoding the A-, B- and C-chain
A;Reference number: S14350; MUID:91174759; PMID:1706597
A;Accession: S14350
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-245 <SEL>
R;Reid, K.B.M.
Biochem. J. 179, 367-371, 1979
A;Title: Complete amino acid sequences of the three collagen-like regions present in su
A;Reference number: A90304; MUID:80020137; PMID:486087
A;Accession: A90304
A;Molecule type: protein
A;Residues: 23-96,'X',98-102,'P',104-130 <RE1>
R;Reid, K.B.M.; Gagnon, J.; Frampton, J.
Biochem. J. 203, 559-569, 1982
A;Title: Completion of the amino acid sequences of the A and B chains of subcomponent C
A;Reference number: A90315; MUID:82283890; PMID:6981411
A;Accession: A90315
A;Molecule type: protein
A;Residues: 131-171,'N',173-177,'W',179-239,'ILPGFSA' <RE2>
C;Comment: The first component of complement is a calcium-dependent complex of the three
ivation of C1r (enzyme), C1s (proenzyme), and the other eight components of complement.
C;Comment: The C1q subcomponent is composed of nine subunits, six of which are disulfid
(see PIR:C1HUQC) chain. Equimolar amounts of the A, B, and C chains are found after red
C;Genetics:
A;Gene: GDB:C1QA
A;Cross-references: GDB:119042; OMIM:120550
A;Map position: 1p36.3-1p34.1
A;Intons: 55/1
C;Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal ho
C;Keywords: complement pathway; glycoprotein; heterodimer; hydroxylysine; hydroxyprolin
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-245/Product: complement subcomponent C1q chain A #status experimental
F;31-109/Domain: collagenous, triple helix #status predicted <COL>
F;116-243/Domain: complement C1q carboxyl-terminal homology <C1Q>
F;26/Disulfide bonds: interchain (to chain B-31) #status experimental
F;33,48,67,100,103/Modified site: 5-hydroxylysine (Lys) #status experimental
F;33,48,67,100,103/Binding site: carbohydrate (Lys) (covalent) #status experimental
F;39,45,54,57,73,85,97/Modified site: 4-hydroxyproline (Pro) #status experimental
F;146/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 19.7%; Score 147; DB 1; Length 245;
Best Local Similarity 30.7%; Pred. No. 1.9e-07;
Matches 43; Conservative 21; Mismatches 72; Indels 4; Gaps 3;

QY 4 TPGPGGSSELRSFAARTTPLEGTSEMAVTFDKYYVNIIGDFDVATGQFRCRVPGAYFFS 63
Db 105 SPGNIKDQPRPAPSAIRNPMPGN--VVFDTVTNQEEPRYNHSGRFVCTVPGYYYFT 162

QY 64 FTAGKAPHKSLSVMLVRNRDEVQALAFDEQRRPGARRAASQASAMLQLDYGDYVWLRHLGA 123
Db 163 FQVLSQWEICLSIVSSSRGQVRRSLGFCDTTNKGLFQVVSQGMVLQLOGQVWVEKDPK 222

QY 124 P-HYALGAPG-ATFSGYLVY 141
Db 223 KGH1YQGSSEADSVFSGFLIF 242

RESULT 14
CGHUID
collagen alpha 1(X) chain precursor - human
N;Alternate names: procollagen alpha 1(X) chain
C;Species: Homo sapiens (man)
C;Date: 22-Nov-1993 #sequence_revision 03-Nov-1995 #text_change 22-Jun-1999
C;Accession: S26396; S30086; S15826; S18249; A43901; I51870; S21856
R;Reichenberger, E.; Beter, F.; LuValle, P.; Olsen, B.R.; von der Mark, K.; Bertling, W
FEBS Lett. 311, 305-310, 1992
A;Title: Genomic organization and full-length cDNA sequence of human collagen X.
A;Reference number: S26396; MUID:93012005; PMID:1397333
A;Accession: S26396
A;Molecule type: DNA

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:09 ; Search time 5.13772 Seconds
(without alignments)
1308.910 Million cell updates/sec

Title: US-10-085-167-2_COPY_17_159
Perfect score: 745
Sequence: 1 LGPTPGGSSSELRSAFSAAR.....PHYALGAPGATFSGYLVYAD 143

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 745 | 100.0 | 329 | 1 CQT4_HUMAN | Q9bxj3 homo sapien |
| 2 | 196 | 26.3 | 285 | 1 CQT2_HUMAN | Q9bxj5 homo sapien |
| 3 | 191 | 25.6 | 244 | 1 APM1_HUMAN | Q15848 homo sapien |
| 4 | 189 | 25.4 | 243 | 1 CQT5_HUMAN | Q9bxj0 homo sapien |
| 5 | 188 | 25.2 | 278 | 1 CQT6_HUMAN | Q9bxj9 homo sapien |
| 6 | 185 | 24.8 | 247 | 1 APM1_MOUSE | Q60994 mus musculu |
| 7 | 181.5 | 24.4 | 258 | 1 C1RF_HUMAN | Q75973 homo sapien |
| 8 | 181.5 | 24.4 | 258 | 1 C1RF_MOUSE | O88992 mus musculu |
| 9 | 175 | 23.5 | 289 | 1 CQT7_HUMAN | Q9bxj2 homo sapien |
| 10 | 170.5 | 22.9 | 281 | 1 CQT1_HUMAN | Q9bxj1 homo sapien |
| 11 | 168.5 | 22.6 | 251 | 1 C1OB_HUMAN | P02746 homo sapien |
| 12 | 167 | 22.4 | 419 | 1 COLE_LEPMA | P98085 lepomis mac |
| 13 | 165.5 | 22.2 | 255 | 1 GLIC_MOUSE | Q9esn4 mus musculu |
| 14 | 163.5 | 21.9 | 246 | 1 CQT3_HUMAN | Q9bxj4 homo sapien |
| 15 | 163.5 | 21.9 | 253 | 1 C1OB_RAT | P31721 rattus norv |
| 16 | 160 | 21.5 | 245 | 1 C1QC_HUMAN | P02747 homo sapien |
| 17 | 156.5 | 21.0 | 508 | 1 OTO1_ONCKE | P83371 oncorhynch |
| 18 | 156 | 20.9 | 245 | 1 C1QA_MOUSE | P98086 mus musculu |
| 19 | 152.5 | 20.5 | 246 | 1 C1QC_MOUSE | Q02105 mus musculu |
| 20 | 151.5 | 20.3 | 253 | 1 C1OB_MOUSE | P14106 mus musculu |
| 21 | 150 | 20.1 | 224 | 1 CERL_MOUSE | P98087 rattus norv |
| 22 | 149 | 20.0 | 193 | 1 CERB_HUMAN | P23435 homo sapien |
| 23 | 149 | 20.0 | 193 | 1 CERB_MOUSE | Q9x171 mus musculu |
| 24 | 148.5 | 19.9 | 674 | 1 CA1A_CHICK | P08125 gallus gall |
| 25 | 148 | 19.9 | 674 | 1 CA1A_BOVIN | P23206 bos taurus |
| 26 | 147 | 19.7 | 245 | 1 C1QA_HUMAN | P02745 homo sapien |
| 27 | 145 | 19.5 | 680 | 1 CA1A_HUMAN | Q03692 homo sapien |
| 28 | 144.5 | 19.4 | 215 | 1 HP27_TAMSI | Q06577 tamias sibi |
| 29 | 144 | 19.3 | 1228 | 1 ECM_HUMAN | Q13201 homo sapien |
| 30 | 141 | 18.9 | 743 | 1 CA18_MOUSE | Q00780 mus musculu |
| 31 | 139.5 | 18.7 | 635 | 1 CA28_HUMAN | P25067 homo sapien |
| 32 | 139.5 | 18.7 | 744 | 1 CA18_RABIT | P14282 oryctolagus |
| 33 | 138.5 | 18.6 | 744 | 1 CA18_HUMAN | P27658 homo sapien |

| | | | | | |
|----|-------|------|-----|--------------|----------------------|
| 34 | 136.5 | 18.3 | 215 | 1 HP25_TAMSI | Q06576 tamias sibi |
| 35 | 135.5 | 18.2 | 170 | 1 CA28_MOUSE | P25318 mus musculu |
| 36 | 132 | 17.7 | 680 | 1 CA1A_MOUSE | Q05306 mus musculu |
| 37 | 126 | 16.9 | 201 | 1 CERL_HUMAN | Q9ntu7 homo sapien |
| 38 | 113.5 | 15.2 | 196 | 1 HP20_TAMSI | Q06575 tamias sibi |
| 39 | 74 | 9.9 | 181 | 1 RUVB_RALSO | Q8Y233 ralonstonia s |
| 40 | 74 | 9.9 | 272 | 1 ASRB_SALTY | P26475 salmonella |
| 41 | 74 | 9.9 | 610 | 1 DNAB_BRECH | Q91cqs brevibacill |
| 42 | 73 | 9.8 | 413 | 1 IF42_NICPL | P41379 nicotiana p |
| 43 | 73 | 9.8 | 789 | 1 PRTF_EBV | P25939 epstein-bar |
| 44 | 69.5 | 9.3 | 440 | 1 PAPA_MOUSE | Q60963 m platelet- |
| 45 | 69 | 9.3 | 355 | 1 IF4X_TOBAC | Q40466 nicotiana t |

ALIGNMENTS

| RESULT 1 | ID | QQT4_HUMAN | STANDARD; | PRT; | 329 AA. |
|-----------------------|--|---|--------------------|-------------------------|---------------------------------------|
| AC | Q9BXJ3; | 28-FEB-2003 (Rel. 41, Created) | | | |
| DT | 28-FEB-2003 (Rel. 41, Last sequence update) | | | | |
| DT | 28-FEB-2003 (Rel. 41, Last annotation update) | | | | |
| DE | Complement-clq tumor necrosis factor-related protein 4 precursor. | | | | |
| GN | C1QTNF4 OR CTRP4. | | | | |
| OS | Homo sapiens (Human). | | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| OX | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | |
| NCBI_TaxID=9606; | | | | | |
| RN | [1] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RA | Holloway J.L., Lok S.; | | | | |
| RT | "Homo sapiens complement-clq tumor necrosis factor-related protein."; | | | | |
| RL | Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases. | | | | |
| CC | -1- SIMILARITY: Contains 2 C1Q domains. | | | | |
| CC | ----- | | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | | | |
| CC | or send an email to license@isb-sib.ch). | | | | |
| CC | ----- | | | | |
| DR | EMBL; AF329838; AAK17962.1; - | | | | |
| DR | Genew; HGNC:14346; C1QTNF4. | | | | |
| DR | InterPro; IPR001073; C1q. | | | | |
| DR | Pfam; PF00386; C1q; 2. | | | | |
| DR | SMART; SM00110; C1Q; 2. | | | | |
| DR | PROSITE; PS01113; C1Q; 2. | | | | |
| KW | Repeat; Signal. | | | | |
| FT | SIGNAL | 1 | 16 | | POTENTIAL. |
| FT | CHAIN | 17 | 329 | | COMPLEMENT-C1Q TUMOR NECROSIS FACTOR- |
| FT | | | | | RELATED PROTEIN 4. |
| FT | DOMAIN | 23 | 159 | | C1Q 1. |
| FT | DOMAIN | 170 | 314 | | C1Q 2. |
| SO | SEQUENCE | 329 AA; | 35265 MW; | 331C7DBF26036915 CRC64; | |
| Query Match | | 100.0%; | Score 745; | DB 1; | Length 329; |
| Best local Similarity | | 100.0%; | Pred. No. 3.4e-70; | | |
| Matches 143; | | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| QY | 1 | LGPTPGGSSSELRSAFSAARTTPLEGTSMAVTFDKVYVNIIGDGFVATGQFRCRVPGAY | 60 | | |
| DB | 17 | LGPTPGGSSSELRSAFSAARTTPLEGTSMAVTFDKVYVNIIGDGFVATGQFRCRVPGAY | 76 | | |
| QY | 61 | FFSFTAGKAPKSLVLMVRNDEVQALAFDEQRRPGARRAASQSAAMLQLDYGDTVWLRL | 120 | | |
| DB | 77 | FFSFTAGKAPKSLVLMVRNDEVQALAFDEQRRPGARRAASQSAAMLQLDYGDTVWLRL | 136 | | |
| QY | 121 | HGAPHYALGAPGATFSGYLVYAD | 143 | | |

Db 137 HGAPHYALGAPGATFSGYLVYAD 159

RESULT 2

COT2_HUMAN STANDARD; PRT; 285 AA.

ID COT2_HUMAN

AC Q9BXJ5;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Complement-clq tumor necrosis factor-related protein 2 precursor.

GN C1QTNF2 OR CTRP2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Piddington C.S., Bishop P.;

RT "Homo sapiens complement-clq tumor necrosis factor-related protein.";

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Muscle;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Bhat N.K., Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Hsieh F., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ueidi T.B., Toshlyuk S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzyzinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -1- SIMILARITY: Contains 1 collagenous domain.

CC -1- SIMILARITY: Contains 1 C1Q domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).

CC -----

DR EMBL; AF329836; AAK17960.1; -.

DR EMBL; BC011699; AAH1699.1; -.

DR Genew; HGNC:14325; C1QTNF2.

DR InterPro; IPR001073; C1Q.

DR InterPro; IPR000087; Collagen.

DR Pfam; PF00386; C1Q; 1.

DR Pfam; PF01391; Collagen; 2.

DR PRINTS; PR00007; COMPLEMENTC1Q.

DR SMART; SM00110; C1Q; 1.

DR PROSITE; PS01113; C1Q; 1.

DR Collagen; Signal.

FT SIGNAL 1 15

FT CHAIN 16 285

FT DOMAIN 40 141

FT DOMAIN 143 285

FT SEQUENCE 285 AA; 29952 MW; 7E31FF9868D4EDFA CRC64;

POTENTIAL.

COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-RELATED PROTEIN 2.

COLLAGEN-LIKE.

C1Q.

Query Match 26.3%; Score 196; DB 1; Length 285;

Best Local Similarity 34.2%; Pred. No. 4.5e-13;

Matches 52; Conservative 30; Mismatches 46; Indels 24; Gaps 8;

QY 5 PGP---GSSELRSAFSA--RTTPLEGTSSEMAVTEFDKVVYVNIIGDFDVAQTGQRCRVPGA 59

Db 138 PGPSCSGSGHTKSAFSAVATKSYPRE--RLPIKFDKILMNEGHYNASGKFCVCGVPGI 194

QY 60 YFFSFTAGKAPPKSLVMLVVRNDEVQALAFDEQRPGARASQSAMLQDYGDVWLR 119

Db 195 YFTYDITLA-NKHLAIGLVHN-GQYRIRTFD--ANTGNHDVASSGTILALKQGDVWVLO 250

QY 120 LHGA-----PHYALGAPGATFSGYLVYAD 143

Db 251 IFYSEQNGLFYDPYWT---DSLFTGLIYAD 278

RESULT 3

APM1_HUMAN STANDARD; PRT; 244 AA.

ID APM1_HUMAN

AC Q15848;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Adiponectin precursor (30 kDa adipocyte complement-related protein) (ACRP30) (Adipose most abundant gene transcript 1) (apm-1) (Gelatin-binding protein).

GN APM1 OR ACRP30 OR GBP28.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Adipose tissue;

RX MEDLINE=96224171; PubMed=8619847;

RA Maeda K., Okubo K., Shimomura I., Funahashi T., Matsuzawa Y., Matsubara K.;

RT "cDNA cloning and expression of a novel adipose specific collagen-like factor, apm1 (Adipose Most abundant Gene transcript 1).";

RL Biochem. Biophys. Res. Commun. 221:286-289(1996).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=99196984; PubMed=10095105;

RA Saito K., Tobe T., Minoshima S., Asakawa S., Sumiya J., Yoda M., Nakano Y., Shimizu N., Tomita M.;

RT "Organization of the gene for gelatin-binding protein (GBP28).";

RL Gene 229:67-73(1999).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=99333693; PubMed=10403784;

RA Schaeffler A., Orso E., Palitzsch K.D., Buechler C., Drobnik W., Fuerst A., Schoelmerich J., Schmitz G.;

RT "The human apm-1, an adipocyte-specific gene linked to the family of TNF's and to genes expressed in activated T cells, is mapped to chromosome 1q21.3-q23, a susceptibility locus identified for familial combined hyperlipidemia (FCH).";

RT Biochem. Biophys. Res. Commun. 260:416-425(1999).

RN [4]

RP CHARACTERIZATION.

RX MEDLINE=20417747; PubMed=10961870;

RA Yokota T., Oritani K., Takahashi I., Ishikawa J., Matsuyama A., Ouchi N., Kihara S., Funahashi T., Tenner A.U., Tomiyama Y., Matsuzawa Y.;

RT "Adiponectin, a new member of the family of soluble defense collagens, negatively regulates the growth of myelomonocytic progenitors and the functions of macrophages.";

RT Blood 96:1723-1732(2000).

RL [5]

RP CHARACTERIZATION.

RX MEDLINE=20440368; PubMed=10982546;

RA Ouchi N., Kihara S., Arita Y., Okamoto Y., Maeda K., Kuriyama H.,

RA Hotta K., Nishida M., Takahashi M., Muraguchi M., Ohmoto Y.,
RA Nakamura T., Yamashita S., Funahashi T., Matsuzawa Y.;
RT "Adiponectin, an adipocyte-derived plasma protein, inhibits
RT endothelial NF-kappaB signaling through a CAMP-dependent pathway.";
RL Circulation 102:1296-1301(2000).
RN [6]
RP FUNCTION.
RX MEDLINE=21372498; PubMed=11479627;
RA Yamauchi T., Kamon J., Waki H., Terauchi Y., Kubota N., Hara K.,
RA Mori Y., Ide T., Murakami K., Tsuboyama-Kasaoka N., Ezaki O.,
RA Akanuma Y., Gavrilova O., Vinson C., Reitman M.L., Kagechika H.,
RA Shudo K., Yoda M., Nakano Y., Tobe K., Nagai R., Kimura S., Tomita M.,
RA Froguel P., Kadowaki T.;
RT "The fat-derived hormone adiponectin reverses insulin resistance
RT associated with both lipodystrophy and obesity.";
RL Nat. Med. 7:941-946(2001).
RN [7]
RP VARIANT ADIPONECTIN DEFICIENCY CYS-112.
RX MEDLINE=20378830; PubMed=10918532;
RA Takahashi M., Arita Y., Yamagata K., Matsukawa Y., Okutomi K.,
RA Horie M., Shimomura I., Hotta K., Kuriyama H., Kihara S., Nakamura T.,
RA Yamashita S., Funahashi T., Matsuzawa Y.;
RT "Genomic structure and mutations in adipose-specific gene,
RT adiponectin.";
RL Int. J. Obes. Relat. Metab. Disord. 24:861-868(2000).
RN [8]
RP VARIANTS ARG-84; MET-117; THR-164; SER-221 AND PRO-241.
RX MEDLINE=21671103; PubMed=11812766;
RA Hara K., Boutin P., Mori Y., Tobe K., Dina C., Yasuda K., Yamauchi T.,
RA Otobe S., Okada T., Eto K., Kadowaki H., Hagura R., Akanuma Y.,
RA Yazaki Y., Nagai R., Taniyama M., Matsubara K., Yoda M., Nakano Y.,
RA Kimura S., Tomita M., Kimura S., Ito C., Froguel P., Kadowaki T.;
RT "Genetic variation in the gene encoding adiponectin is associated with
RT an increased risk of type 2 diabetes in the Japanese population.";
RL Diabetes 51:536-540(2002).
CC -!- FUNCTION: IMPORTANT NEGATIVE REGULATOR IN HEMATOPOIESIS AND IMMUNE
CC SYSTEMS; MAY BE INVOLVED IN ENDING INFLAMMATORY RESPONSES THROUGH
CC ITS INHIBITORY FUNCTIONS. INHIBITS ENDOTHELIAL NF-KAPPAB SIGNALING
CC THROUGH A CAMP-DEPENDENT PATHWAY. INHIBITS TNF-ALPHA-INDUCED
CC EXPRESSION OF ENDOTHELIAL ADHESION MOLECULES. INVOLVED IN THE
CC CONTROL OF FAT METABOLISM AND INSULIN SENSITIVITY.
CC -!- SUBUNIT: HOMODIGOMER (POTENTIAL).
CC -!- TISSUE SPECIFICITY: Secreted.
CC -!- secreted into plasma.
CC -!- DISEASE: Defects in APM1 are the cause of adiponectin deficiency
CC [MIM:605441], resulting in very low concentration of plasma
CC adiponectin. Decreased adiponectin plasma levels are associated
CC with obesity insulin resistance, and diabetes type 2.
CC -!- PHARMACEUTICAL: Adiponectin might be used in the treatment of
CC diabetes type 2 and insulin resistance.
CC -!- SIMILARITY: Contains 1 collagenous domain.
CC -!- SIMILARITY: Contains 1 C1q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D45371; BAA08227.1; -.
DR EMBL; AB012165; BAA86716.1; -.
DR EMBL; AB012164; BAA86716.1; JOINED.
DR EMBL; AJ131460; CAB52413.1; -.
DR EMBL; AJ131461; CAB52413.1; JOINED.
DR PIR; JC4708; JC4708.
DR MIM; 605441; -.
DR GO; GO:0006091; P:energy pathways; TAS.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.

DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR ProDom; PD000007; C1q_helix; 1.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Hormone; Collagen; Signal; Repeat; Hydroxylation; Plasma;
KW Polymorphism; Disease mutation; Obesity; Diabetes mellitus.
FT SIGNAL 1 14
FT CHAIN 15 244
FT DOMAIN 42 107
FT DOMAIN 108 244
FT DISULFID 36 36
FT MOD_RES 44 44
FT MOD_RES 47 47
FT MOD_RES 53 53
FT MOD_RES 62 62
FT MOD_RES 71 71
FT MOD_RES 76 76
FT MOD_RES 86 86
FT MOD_RES 95 95
FT MOD_RES 104 104
FT VARIANT 84 84
FT VARIANT 112 112
FT VARIANT 117 117
FT VARIANT 164 164
FT VARIANT 221 221
FT VARIANT 241 241
SQ SEQUENCE 244 AA; 26414 MW; 64D8C6C1204B1018 CRC64;
Query Match 25.6%; Score 191; DB 1; Length 244;
Best Local Similarity 31.8%; Pred. No. 1.2e-12;
Matches 47; Conservative 29; Mismatches 54; Indels 18; Gaps 6;
QY 5 PGPSSLSRSPSAARTPLRGTSMAVTFPKYVNVNIGDFDVATGQRCRVPGYFFSF 64
Db 104 PGEAYVYRSASFVGLFTYVT-IPNMPFRFKIFYNQNHYDGSSTGKHCNIPGLYFAY 162
QY 65 TAGKAPH-----KSLVMLVRNRDEVQALAFDEQRPGARRAASQSANLQLDYGDVWLR 119
Db 163 -----HITVYMKDVKVSLEFK-KDKAMLFYDQYQENNVQ-ASGSVLLHLEVGDPQVWLQ 214
QY 120 LHGAPH----YALGAPGATFSGLYVAD 143
Db 215 VYGEGERNGLYADNDNDSTFTGFLLYHD 242
RESULT 4
CQTS HUMAN STANDARD; PRT; 243 AA.
AC Q9BXJ0; Q9UFY4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Complement-clq tumor necrosis factor-related protein 5 precursor.
GN C1QTNF5 OR CTRP5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Sheppard P.O., Humes J.M.;
RT "Homo sapiens complement-clq tumor necrosis factor-related protein.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 25-243 FROM N.A.
RC TISSUE=Uterus;

RA Ottenweelder B., Obermaier B., Mewes H.-W., Gaassenhuber J.,
RA Wiemann S.;
RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C1Q domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
DR EMBL; AF329841; AAK17965.1; -
DR EMBL; AL110261; CAB53702.1; -
DR PIR; T14782; T14782.
DR Genew; HGNC:14344; C1QTNF5.
DR InterPro; IPR001073; C1Q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; FALSE_NEG.
KW Collagen; Signal.
FT SIGNAL 1 15
FT CHAIN 16 243
FT FT 16 243
FT FT 30 95
FT DOMAIN 97 243
SQ SEQUENCE 243 AA; 25298 MW; 7CCDA65CDA7EB784 CRC64;
COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-RELATED PROTEIN 5.
COLLAGEN-LIKE.

Query Match 25.4%; Score 189; DB 1; Length 243;
Best Local Similarity 35.3%; Pred. No. 2e-12;
Matches 53; Conservative 28; Mismatches 57; Indels 12; Gaps 6;

QY 2 GPTPGSSSEL--RSAFSAART-TPLEGTSEMAVTFDKVYVNIQGDVATGQRCRVP 58
DB 90 GPTPGAGECVPRPSAFSAKRSSESRVPPSDALPDRVLVNEQGHYDAVTGKTCQVPG 149
QY 59 AYPFSFTAGKAPHKSLSVMLVRNDEVQAL--AFDEQRPGARRAASQAMQLDYGDYV 116
DB 150 VYFAVHA-TVYRASLQFDLVKNGESIASFQFGGWPMPA---SLSGAMVRLPEPDQV 205
QY 117 WLRHGAAPH---YALGAPGATFGGLVYAD 143
DB 206 WVQGVGVGYIGIYASIKTSTFGFLVYSD 235

RESULT 5
CQT6 HUMAN STANDARD; PRT; 278 AA.
ID CQT6_HUMAN
AC Q9BXI9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement-clq tumor necrosis factor-related protein 6 precursor.
GN C1QTNF6 OR CTRP6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Piddington C.S., Sheppard P.O.;
RT "Homo sapiens complement-clq tumor necrosis factor-related protein.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C1Q domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
DR EMBL; AF329842; AAK17966.1; -
DR EMBL; BC020551; AAH20551.1; -
DR Genew; HGNC:14343; C1QTNF6.
DR InterPro; IPR001073; C1Q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR PROSITE; PS01113; C1Q; FALSE_NEG.
KW Collagen; Signal.
FT SIGNAL 1 46
FT CHAIN 47 278
FT FT 47 278
FT FT 97 138
FT DOMAIN 139 278
FT CARBOHYD 91 91
FT CONFLICT 21 21
SQ SEQUENCE 278 AA; 30861 MW; 27A82CA863F23D47 CRC64;

Query Match 25.2%; Score 188; DB 1; Length 278;
Best Local Similarity 35.5%; Pred. No. 3e-12;
Matches 59; Conservative 18; Mismatches 61; Indels 28; Gaps 7;

QY 2 GP--TPPGSS-----ELRS-----AFSAARTPL-EGTSEMAVTFDKVYVNIQ 43
DB 115 GPQGEPPQGSXGDKGEMSGPAPCQKRFFAFSVGRKTAHSGEDFQTLLEFRVFNLDG 174
QY 44 DFDVATGQFCRVRPGAYFFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRPGARR-AA 102
DB 175 CFDMATGQFAAPLRGIYFFSLNVHSMWNYKETYVHIMNQEAVIL---YAQPSERSIMQ 230
QY 103 SQSAMQLDYGDYVWLR-----HGAPHYALGAPGATFGGLVYAD 143
DB 231 SQSVMLDLAYGDRVWRLFKQRENAIYSNDFTYITFGHLIKAE 276

RESULT 6
APM1 MOUSE STANDARD; PRT; 247 AA.
ID APM1_MOUSE
AC Q60994; Q62400; Q9DC68;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Adiponectin precursor (30 kDa adipocyte complement-related protein)
DE (ACRP30) (adipocyte specific protein AdipoQ).
GN APM1 OR ACRP30 OR ADIPOQ.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adipocyte;
RX MEDLINE=96070757; PubMed=7592907;
RA Scherer P.E., Williams S., Fogliano M., Baldini G., Lodish H.F.;
RT "A novel serum protein similar to C1q, produced exclusively in
RT adipocytes.";
RL J. Biol. Chem. 270:26746-26749(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fibroblast;
RX MEDLINE=96209999; PubMed=8631877;
RA Hu E., Liang P., Spiegelman B.M.;
RT "AdipoQ is a novel adipose-specific gene dysregulated in obesity.";
RL J. Biol. Chem. 271:10697-10703(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX PubMed=11162643;
RA Das K., Lin Y., Widen E., Zhang Y., Scherer P.E.;
RT "Chromosomal localization, expression pattern, and promoter analysis
RT of the mouse gene encoding adipocyte-specific secretory protein
RT Acrp30.";
RL Biochem. Biophys. Res. Commun. 280:1120-1129(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guelincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weltz C., Wittaker C., Wilming L.,
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [5]
RP FUNCTION.
RX MEDLINE=21372498; PubMed=11479627;
RA Yamauchi T., Kamon J., Waki H., Terachi Y., Kubota N., Hara K.,
RA Mori Y., Ide T., Murakami K., Tsuboyama-Kasaoka N., Ezaki O.,
RA Akanuma Y., Gavrilova O., Vinson C., Reitman M.L., Kagechika H.,
RA Shudo K., Yoda M., Nakano Y., Tobe K., Nagai R., Kimura S., Tomita M.,
RA Froguel P., Kadowaki T.;
RT "The fat-derived hormone adiponectin reverses insulin resistance
RT associated with both lipatrophy and obesity.";
RL Nat. Med. 7:941-946(2001).
RN [6]
RP FUNCTION.
RX MEDLINE=21372499; PubMed=11479628;
RA Berg A.H., Combs T.P., Du X., Brownlee M., Scherer P.E.;
RT "The adipocyte-secreted protein Acrp30 enhances hepatic insulin
RT action.";
RL Nat. Med. 7:947-953(2001).
CC -!- FUNCTION: IMPORTANT NEGATIVE REGULATOR IN HEMATOPOIESIS AND IMMUNE

CC SYSTEMS; MAY BE INVOLVED IN ENDING INFLAMMATORY RESPONSES THROUGH
CC ITS INHIBITORY FUNCTIONS. INHIBITS ENDOTHELIAL NF-KAPPAB SIGNALING
CC THROUGH A CAMP-DEPENDENT PATHWAY. INHIBITS TNF-ALPHA-INDUCED
CC EXPRESSION OF ENDOTHELIAL ADHESION MOLECULES. INVOLVED IN THE
CC CONTROL OF FAT METABOLISM AND INSULIN SENSITIVITY.
CC -!- SUBUNIT: Homooligomer.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Synthesized exclusively by adipocytes and
CC secreted into plasma.
CC -!- INDUCTION: DURING HORMONE-INDUCED ADIPOSE DIFFERENTIATION AND
CC ACTIVATED BY INSULIN.
CC -!- SIMILARITY: Contains 1 collagenous domain.
CC -!- SIMILARITY: Contains 1 C1q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U37222; AAA0543.1; -.
DR EMBL; U49915; AAB06706.1; -.
DR EMBL; AF304466; AAK13417.1; -.
DR EMBL; AK003138; BAB22597.1; -.
DR PDB; 1C28; 07-SEP-99.
DR MGD; MGI:106675; Acrp30.
DR GO; GO:0005576; C:extracellular; IDA.
DR GO; GO:0005515; F:protein binding activity; IPI.
DR GO; GO:0006635; P:fatty acid beta-oxidation; IMP.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR ProDom; PD000007; C1q_helix; 1.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Hormone; Collagen; Signal; Repeat; Hydroxylation; Plasma;
KW Polymorphism; 3D-structure.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 247 ADIPONECTIN.
FT DOMAIN 45 110 COLLAGEN-LIKE.
FT DOMAIN 111 247 C1Q.
FT DISULFID 39 39 INTERCHAIN (BY SIMILARITY).
FT MOD_RES 47 47 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 50 50 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 56 56 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 65 65 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 79 79 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 98 98 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 107 107 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 113 113 M -> V.
FT CONFLICT 50 50 P -> S (IN REF. 2).
FT CONFLICT 74 74 A -> S (IN REF. 2).
FT CONFLICT 117 117 A -> G (IN REF. 2).
FT CONFLICT 148 148 G -> N (IN REF. 2).
FT CONFLICT 243 243 Y -> F (IN REF. 2).
SQ SEQUENCE 247 AA; 26841 MW; 137B687D873988C4 CRC64;

Query Match 24.8%; Score 185; DB 1; Length 247;
Best Local Similarity 31.8%; Pred. No. 5.3e-12;
Matches 47; Conservative 29; Mismatches 54; Indels 18; Gaps 7;

OY 5 PGPSSSELSAFAAARTTPLEGTSSEMAVTFDKVYVNIIGGDFDVATGQFCRCRVGAYFFSF 64
DB 107 PGEAAAYRSAFVSGLVLETRVT-VENVPIRFTKIFYNQNHYDGSFGKFCYCNIPGLYFSY 165
OY 65 TAKGAPH-----KSLSVMLVRNRDEVQALAFDEQRRPGARRASQASAMLQLDYGDYVWLR 119
DB 166 -----HITVYMKDVKVSFLK-KDKAVLFTYDQYQEKVNDQ-ASGSVLLHLEVGQVWLQ 217

QY 120 LHG-APH--YALGAPGATGSGILVYAD 143
Db 218 VYGDGDHNGLYADNVNDSTFTGFLLYHD 245

RESULT 7
CIRF_HUMAN STANDARD; PRT; 258 AA.
ID CIRF_HUMAN 075973;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE C1q-related factor precursor.
GN C1QRF OR CRF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99097006; PubMed=9878755;
RA Berube N.G., Swanson X.H., Berttram M.J., Kittle J.D., Didenko V.,
RA Baskin D.S., Smith J.R., Pereira-Smith O.M.;
RT "Cloning and characterization of CRF, a novel C1q-related factor,
RT expressed in areas of the brain involved in motor function.";
RL Brain Res. Mol. Brain Res. 63:233-240(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA Nickerson D.A.;
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAINSTEM.
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C1Q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF095154; AAC64186.1; -.
DR EMBL; AF410771; AAK95248.1; -.
DR EMBL; BC008798; AAH08798.1; -.
DR GO; GO:0007626; P:locomotory behavior; NAS.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.

DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Collagen; Signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 258 C1Q-RELATED FACTOR.
FT DOMAIN 67 115 COLLAGEN-LIKE.
FT DOMAIN 123 258 C1Q.
SQ SEQUENCE 258 AA; 26452 MW; 52C51CDF59CAE2BF CRC64;
Query Match 24.4%; Score 181.5; DB 1; Length 258;
Best Local Similarity 32.1%; Pred. No. 1.3e-11;
Matches 54; Conservative 22; Mismatches 51; Indels 41; Gaps 8;
QY 2 GP--TPGPSSEL-----RSAFSARTPELGTSSEMAVTFDKVYVYVIGGDFDVAT 49
Db 106 GPPGLPGAGSGAISTATYTTVPRAVAFYAGLKNPHEGYE--VLKFDVVTNIGNNYDAAS 163
QY 50 GQFRCRVPGAYPFPSFTAGKAPHKSLSVMLVRNDEVOALAFDEQRRPGARRA----- 101
Db 164 GKFTCNIPGTFFFTY-----H-----VLMRGDGTSMWA--DLCKNGQVRASAIQDAD 210
QY 102 -----ASQSAMLQDYCDTWLRLHGA-PHYALGAPGATGSGILVYAD 143
Db 211 QNYDYASNSVILHLDADEVFIKLDGKAHGNSNKYSTPSGFIITSD 258

RESULT 8
CIRF_MOUSE STANDARD; PRT; 258 AA.
ID CIRF_MOUSE 088952;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C1q-related factor precursor.
GN C1QRF OR CRF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99097006; PubMed=9878755;
RA Berube N.G., Swanson X.H., Berttram M.J., Kittle J.D., Didenko V.,
RA Baskin D.S., Smith J.R., Pereira-Smith O.M.;
RT "Cloning and characterization of CRF, a novel C1q-related factor,
RT expressed in areas of the brain involved in motor function.";
RL Brain Res. Mol. Brain Res. 63:233-240(1999).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAINSTEM. MORE ABUNDANT IN AREAS
CC OF THE NERVOUS SYSTEM INVOLVED IN MOTOR FUNCTION, SUCH AS THE
CC PURKINJE CELLS OF THE CEREBELLUM, THE ACCESSORY OLIVARY NUCLEUS,
CC THE PONS AND THE RED NUCLEUS.
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C1Q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF095155; AAC64187.1; -.
DR MGD; MGI:1344400; C1qrf.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.

DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Collagen; Signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 258 C1Q-RELATED FACTOR.
FT DOMAIN 67 115 COLLAGEN-LIKE.
FT DOMAIN 123 258 C1Q.
SQ SEQUENCE 258 AA; 26485 MW; F776E2D206EBF763 CRC64;

Query Match 24.4%; Score 181.5; DB 1; Length 258;
Best Local Similarity 32.1%; Pred. No. 1.3e-11;
Matches 54; Conservative 22; Mismatches 51; Indels 41; Gaps 8;

QY 2 GP--TPGSGSEL-----RSAFSAARTTPLEGTSEMAVTFDKYVNIIGDFDVAT 49
DB 106 GPPGLPGSGSGAISTATYTPVPRVAFYAGLKNPHEGYE--VLKFDVVTNLGNNDYDAAS 163
QY 50 GQFRCRVPGAYFFSFTAGKAPHKSLSVMLVNRDEVQALAFDEQRRPGARRA----- 101
DB 164 GKFTCNIPGTFFTY-----H-----VLNRGGDGTSMWA--DLCKNGQVRASAIQDAD 210
QY 102 -----ASQSAMLQLDYGDVTWLRHLGA-PHYALGAPGATFSGYLVYAD 143
DB 211 QNYDYASNSVILHLDAGDEVFIKLDGKAHGNSNKYSTPFGFIYSD 258

RESULT 9
CQT7 HUMAN STANDARD; PRT; 289 AA.
AC Q9BXJ2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement-c1q tumor necrosis factor-related protein 7 precursor.
GN C1QTNF7 OR CTRP7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Piddington C.S., Sheppard P.O., Bishop P., Lasser G.W.;
RT "Homo sapiens complement-c1q tumor necrosis factor-related protein.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Wortley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- SIMILARITY: Contains 1 collagenous domain.
CC -!- SIMILARITY: Contains 1 C1Q domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF329839; AAK17963.1; -
DR EMBL; BC022187; AAH22187.1; -
DR Genew; HGNC:14342; C1QTNF7.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 2.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Collagen; Signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 289 COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT DOMAIN 38 139 RELATED PROTEIN 7.
FT DOMAIN 141 276 COLLAGEN-LIKE.
SQ SEQUENCE 289 AA; 30683 MW; A61609FF86D26946 CRC64;

Query Match 23.5%; Score 175; DB 1; Length 289;
Best Local Similarity 33.5%; Pred. No. 7e-11;
Matches 56; Conservative 21; Mismatches 58; Indels 32; Gaps 8;

QY 1 LGTPPGP-----GSSELSAFSAARTT--PLEGTSEMAVTFDKYVNI 40
DB 118 IGP-PGRKDRGEQDGLPGVCRGSIVLKSAFSVGITTSYEE--RLPIIFNKVLFN 173
QY 41 IGGDFDVATGQFRCRVPGAYFFSFTAGKAPHKSLSVMLVNRDEVQALAFDEQRRPGARR 100
DB 174 EGEHYNPATGKFCIFAPGFIYFSYDITLA-NKHLAIGLVHN-GQYRIKTFD--ANTGNHD 229
QY 101 AASQSAMLQLDYGDVTWLRHLGAPHYAL---GAPGATFSGYLVYAD 143
DB 230 VASGSTVIYIQLPEDEWVLEIFFTDQNGLFSDPGWADSLFSGFLLYVD 276

RESULT 10
CQT1 HUMAN STANDARD; PRT; 281 AA.
AC Q9BXJ1; Q96NF2; Q9GZR4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement-c1q tumor necrosis factor-related protein 1 precursor
(G protein coupled receptor interacting protein) (GIP).
GN C1QTNF1 OR CTRP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Innamorati G., Le Guill C., Whang I., Birnbaumer M.;
RT "GIP, a putative GPCR interacting protein.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Sheppard P.O.;
RT "Homo sapiens complement-c1q tumor necrosis factor-related protein.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 83-281 FROM N.A.
RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA Hotta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J.,
RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishi S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuo Y., Nagai K., Isogai T.,
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C1Q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ272138; CAC20425.1; ALT_INIT.
DR EMBL; AF232905; AAG44303.1; ALT_INIT.
DR EMBL; AF329840; AAK17964.1; -
DR EMBL; BC021553; AAH21553.1; -
DR EMBL; AK055541; BAB70947.1; -
DR Genew; HGNC:14324; C1QTNF1.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; FALSE_NEG.
KW Collagen; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 281 COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT DOMAIN 99 140 RELATED PROTEIN 1.
FT DOMAIN 147 281 C1Q.
FT CONFLICT 22 22 L -> P (IN REF. 1).
FT CONFLICT 241 241 R -> Q (IN REF. 1).
SQ SEQUENCE 281 AA; 31743 MW; 49E248CB88ACFB7C CRC64;
Query Match 22.9%; Score 170.5; DB 1; Length 281;
Best local Similarity 35.7%; Pred. No. 2e-10;
Matches 51; Conservative 18; Mismatches 63; Indels 11; Gaps 5;
OY 5 PGPGSSSELRSAFSAARTPLEGTS-EMAVTFDKVYVNIIGGDFDVATGQRCRVPGAYFFS 63
DB 137 PGERCKSHYAAFSVGRKKPMHSNHYQYVTFDTEFVNLVDHFMFTGKRYCYVPGLYFFS 196
OY 64 FTAGKAPHKSLSVMLVRNDEVOALAFDEQRRPGARR-AASQASAMQLDYGDTVWLRLH- 121
DB 197 LNVHTWNOKETVYHIMKNEEVILF---AQVGDRSIMOSQSLMLELRDQDVWVRLYK 252

OY 122 GAPHYALGAPG---ATPSGYLV 140
DB 253 GERENAIIFSELDITYITFSGYLV 275
RESULT 11
C1QB_HUMAN STANDARD; PRT; 251 AA.
ID C1QB_HUMAN
AC P02746; O96H17;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement C1q subcomponent, B chain precursor.
GN C1QB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86076906; PubMed=3000358;
RA Reid K.B.M.;
RT "Molecular cloning and characterization of the complementary DNA and
RT gene coding for the B-chain of subcomponent C1q of the human
RT complement system.";
RL Biochem. J. 231:729-735(1985).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 26-133.
RX MEDLINE=80020137; PubMed=486087;
RA Reid K.B.M.;
RT "Complete amino acid sequences of the three collagen-like regions
RT present in subcomponent C1q of the first component of human
RT complement.";
RL Biochem. J. 179:367-371(1979).
RN [4]
RP SEQUENCE OF 26-193.
RX MEDLINE=79041552; PubMed=708376;
RA Reid K.B.M., Thompson E.O.P.;
RT "Amino acid sequence of the N-terminal 108 amino acid residues of the
RT B chain of subcomponent C1q of the first component of human
RT complement.";
RL Biochem. J. 173:863-868(1978).
RN [5]
RP SEQUENCE OF 134-251.
RX MEDLINE=82283890; PubMed=6981411;
RA Reid K.B.M., Gagnon J., Frampton J.;
RT "Completion of the amino acid sequences of the A and B chains of
RT subcomponent C1q of the first component of human complement.";
RL Biochem. J. 203:559-569(1982).

RN [6]
RP SEQUENCE OF 224-251 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=85038855; PubMed=6208566;
RA Reid K.B.M., Bentley D.R., Wood K.J.;
RT "Cloning and characterization of the complementary DNA for the B
chain of normal human serum C1q";
RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 306:345-354(1984).
RN [7]
RP REVIEW OF C1Q DEFICIENCY.
RX MEDLINE=98450587; PubMed=9777412;
RA Petry F.;
RT "Molecular basis of hereditary C1q deficiency.";
RL Immunobiology 199:286-294(1998).
CC -!- FUNCTION: C1Q ASSOCIATES WITH THE PROENZYMES C1R AND C1S TO YIELD
C1, THE FIRST COMPONENT OF THE SERUM COMPLEMENT SYSTEM. THE
COLLAGEN-LIKE REGIONS OF C1Q INTERACT WITH THE CA(2+)-DEPENDENT
C1R(2)C1S(2) PROENZYME COMPLEX, AND EFFICIENT ACTIVATION OF C1
TAKES PLACE ON INTERACTION OF THE GLOBULAR HEADS OF C1Q WITH THE
FC REGIONS OF IGG OR IGM ANTIBODY PRESENT IN IMMUNE COMPLEXES.
CC -!- SUBUNIT: C1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q, R
AND S IN THE MOLAR RATION OF 1:2:2. C1Q SUBCOMPONENT IS COMPOSED
OF NINE SUBUNITS, SIX OF WHICH ARE DISULFIDE-LINKED DIMERS OF THE
A AND B CHAINS, AND THREE OF WHICH ARE DISULFIDE-LINKED DIMERS OF
THE C CHAIN.
CC -!- PTM: O-linked glycans consist of Glc-Gal disaccharides bound to
the oxygen atom of post-translationally added hydroxyl groups.
CC -!- DISEASE: Defects in C1QB are a cause of C1q deficiency
[MIM:120570]. It is a rare genetic disorder which is associated
with recurrent infections and a high prevalence of lupus
erythematosus-like symptoms. It is characterized by a loss of
activation of the complement classical pathway.
CC -!- SIMILARITY: Contains 1 collagenous domain.
CC -!- SIMILARITY: Contains 1 C1q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X03084; CAA26880.1; -.
DR EMBL; BC008983; AAH08983.1; ALT_INIT.
DR EMBL; M36278; AAC41692.1; -.
DR Genew; HGNC:1242; C1QB.
DR MIM; 120570; -.
DR GO; GO:0005602; C:complement component C1q complex; TAS.
DR GO; GO:0003811; F:complement activity; TAS.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Complement pathway; Plasma; Hydroxylation; Glycoprotein; Collagen;
KW Repeat; Signal; Disease mutation; Pyrrolidone carboxylic acid.
FT SIGNAL 1 25
FT CHAIN 26 251
FT DOMAIN 29 112
FT DOMAIN 113 251
FT MOD_RES 26 26
FT DISULFID 29 29
FT MOD_RES 33 33
FT MOD_RES 36 36
FT MOD_RES 39 39
FT MOD_RES 42 42
FT MOD_RES 51 51
FT MOD_RES 54 54
FT MOD_RES 57 57
FT CARBOHYD 57 57
O-LINKED (GAL. . .).

FT MOD_RES 60 60
FT CARBOHYD 60 60
FT MOD_RES 63 63
FT MOD_RES 75 75
FT MOD_RES 81 81
FT MOD_RES 84 84
FT MOD_RES 90 90
FT MOD_RES 96 96
FT CARBOHYD 96 96
FT MOD_RES 99 99
FT MOD_RES 102 102
FT MOD_RES 105 105
FT MOD_RES 108 108
FT CARBOHYD 108 108
FT VARIANT 40 40
FT CONFLICT 26 26
FT CONFLICT 83 83
FT CONFLICT 98 98
SQ SEQUENCE 251 AA; 26459 MW; 78C5752E267A0EF7 CRC64;
Query Match 22.6%; Score 168.5; DB 1; Length 251;
Best Local Similarity 27.7%; Pred. No. 2.8e-10;
Matches 44; Conservative 34; Mismatches 62; Indels 19; Gaps 4;
QY 1 LGPTPGPS-----SELRSASFAARTPLEGTSEMAVTFDKYVYVNGGDFD 46
Db 93 MGPKGGPAGPAGPCKGESGDKATOKIAFSATRTINVPRLRDQTRFDVITMNNNVE 152
QY 47 VATGQFCRVPVPGAYPFPSFTAGKAPHKSLVNLVNRDEVQALAFDEQRRPGARRAASQSA 106
Db 153 PRSGKFTCKVPGLYFTYTHA--SSRGNLCVNLMGKGERAOKVTFCDIAYVTFQVTTGGM 210
QY 107 MLQLDYGDTVWLRLHGAPHYALGAPGAT--FSGYLVYAD 143
Db 211 VLKLEGGENVFLQATD-KNSLLGMEGANSIFSGLLPD 248
RESULT 12
COLLE LEPMA
ID COLLE LEPMA STANDARD; PRT; 419 AA.
AC P98085; Q91080;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Inner ear-specific collagen precursor (Saccular collagen).
OS Lepomis macrochirus (Bluegill).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Centrarchidae; Lepomis.
OX NCBI_TaxID=13106;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95167486; PubMed=7863331;
RA Davis J.G., Oberholzer J.C., Burns F.R., Greene M.I.;
RT "Molecular cloning and characterization of an inner ear-specific
structural protein";
RL Science 267:1031-1034(1995).
RN [2]
RP CONCEPTUAL TRANSLATION.
RA Gibson T.;
RL Submitted (MAR-1995) to the SWISS-PROT data bank.
CC -!- FUNCTION: FORMS A MICROSTRUCTURAL MATRIX WITHIN THE OTOLITHIC
MEMBRANE (PROBABLY).
CC -!- TISSUE SPECIFICITY: SPECIALIZED SECRETORY SUPPORTING CELLS AT THE
OUTER PERIMETER OF THE SACCULAR EPITHELIUM.
CC -!- SIMILARITY: Contains 1 C1Q domain.
CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A FRAMESHIFT HAD TO BE
INTRODUCED FOR POSITIONS 391-419 SO AS TO MAXIMIZE THE SIMILARITY
WITH OTHER SHORT-CHAIN COLLAGENS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; U17431; AAA69978.1; ALT_FRAME.

DR InterPro; IPR001073; C1q.

DR InterPro; IPR000087; Collagen.

DR Pfam; PF00386; C1q; 1.

DR Pfam; PF01391; Collagen; 3.

DR PRINTS; PR00007; COMPLEMENTC1Q.

DR PRODOM; PD000007; C1q_helix; 2.

DR SMART; SM00110; C1Q; 1.

DR PROSITE; PS01113; C1Q; 1.

KW Extracellular matrix; Repeat; Collagen; Signal.

FT SIGNAL 1 19 POTENTIAL.

FT CHAIN 20 419 INNER EAR-SPECIFIC COLLAGEN.

FT DOMAIN 20 57 NONHELICAL REGION (NC2).

FT DOMAIN 58 274 TRIPLE-HELICAL REGION (COL1).

FT DOMAIN 275 419 NONHELICAL REGION (NC1).

FT DOMAIN 272 419 C1Q.

FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 419 AA; 43634 MW; 570CDB9675FC0F39 CRC64;

Query Match 22.4%; Score 167; DB 1; Length 419;

Best Local Similarity 32.9%; Pred. No. 7.3e-10;

Matches 52; Conservative 20; Mismatches 56; Indels 30; Gaps 8;

OY 2 GPTPGPGSS--ELRSAFSA---ARTPLEGTSEMAVTFDKVYVNIIGDQFVATGQFRCR 55

DB 266 GPKGPPGESVEQIRSAFSVGLFPSRSFP---PPLPVKFDKVFYNGEGHWDPTLNKFNVT 322

OY 56 VPGAYFFSF--TAGKAPHKSL---SVMLVRNDEVQALAFDEQRRPGARRASQSAVLQ 109

DB 323 YPGVYLFYSYHITVRNRPVRAALVNVGVKRLTRDSLXGQDIDQ-----ASNLAHLH 373

OY 110 LDYGDVTWLR---HGAPHYALGAPATFSGYLVYAD 143

DB 374 LTDGDQVWLETRLDWNGX--YSSSEDDSTFSGFLLYPD 409

RESULT 13

GLIC_MOUSE STANDARD; PRT; 255 AA.

ID GLIC_MOUSE STANDARD; PRT; 255 AA.

AC Q9ESN4; 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Gliacolin precursor (C1q-like protein).

GN C1QL.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20428709; PubMed=10862616;

RA Koide T., Aso A., Yorihuzi T., Nagata K.;

RT "Conformational requirements of collagenous peptides for recognition

RT by the chaperone protein HSP47."

RL J. Biol. Chem. 275:27957-27963(2000).

CC -1- TISSUE SPECIFICITY: EXPRESSED IN GLIAL CELLS.

CC -1- SIMILARITY: Contains 1 collagenous domain.

CC -1- SIMILARITY: Contains 1 C1q domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AB044560; BAB15806.1; -.

DR MGD; MG1:2387350; C1q1.

DR GO; GO:0005515; F:protein binding activity; IPI.

DR InterPro; IPR001073; C1q.

DR InterPro; IPR000087; Collagen.

DR Pfam; PF00386; C1q; 1.

DR Pfam; PF01391; Collagen; 1.

DR PRINTS; PR00007; COMPLEMENTC1Q.

DR SMART; SM00110; C1Q; 1.

DR PROSITE; PS01113; C1Q; 1.

KW Collagen; Signal.

FT SIGNAL 1 20 POTENTIAL.

FT CHAIN 21 255 GLIACOLIN.

FT DOMAIN 61 111 COLLAGEN-LIKE.

FT DOMAIN 120 255 C1Q.

SQ SEQUENCE 255 AA; 26687 MW; 529FBAF4B2191BC1 CRC64;

Query Match 22.2%; Score 165.5; DB 1; Length 255;

Best Local Similarity 30.4%; Pred. No. 5.9e-10;

Matches 48; Conservative 23; Mismatches 66; Indels 21; Gaps 5;

OY 2 GPTPGPG-----SSELSAFSAARTPLEGTSEMAVTFDKVYVNIIGDQFVAT 49

DB 103 GPPGAPGLNAGAIASATYSTVPKIAFYAGLKRQHEGYE--VLKEDDVNTLGNHYDPTT 160

OY 50 GQFRCRVPGAYFFSF--TAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRASQSA 106

DB 161 GKFTCSIPGIYFTYHVLNRGGDGTSMWADLCKNNQVRASAIQADADQN--YDYASNSV 217

OY 107 MLOLDYGDVTWLRHGA-PHYALGAPATFSGYLVYAD 143

DB 218 VLHLEPGDEVYIKLDGKAHGNNKYSTFSGFILYAD 255

RESULT 14

COT3_HUMAN

ID COT3_HUMAN STANDARD; PRT; 246 AA.

AC Q9BXJ4; Q96KY1; 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Complement-c1q tumor necrosis factor-related protein 3 precursor

DE (Secretory protein COR326).

GN C1QTNF3 OR CTRP3.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Piddington C.S., Bishop P.;

RT "Homo sapiens complement-c1q tumor necrosis factor-related protein.";

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Maeda T., Hayashi A., Saito T.;

RT "Molecular cloning, chromosomal localization, and genomic structure of

RT the human COR326 gene.";

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 121-246 FROM N.A.

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton A., Ketteman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.T., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences";
CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -I- SIMILARITY: Contains 1 collagenous domain.
CC -I- SIMILARITY: Contains 1 C1Q domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outpost at
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```
DR   EMBL; AF329837; AAK17961.1; -.
DR   EMBL; AF326976; AAK70344.1; -.
DR   EMBL; BC016021; AAH16021.1; -.
DR   Genew; HGNC:14326; C1QTNF3.
DR   InterPro; IPR001073; C1Q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 1.
DR   SMART; SM00110; C1Q; 1.
DR   PROSITE; PS01113; C1Q; 1.
KW   Collagen; Signal.
FT   SIGNAL      1          22    POTENTIAL.
FT   CHAIN       23         246   COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT                                     RELATED PROTEIN 3.
FT               51         113   COLLAGEN-LIKE.
FT   DOMAIN      114        246   C1Q.
FT   CONFLICT    214        214   K->E (IN REF. 3).
SQ   SEQUENCE    246 AA; 26994 MW; C589B6C3A73E5D29 CRC64;
```

Query Match 21.9%; Score 163.5; DB 1; Length 246;
Best Local Similarity 35.2%; Pred. No. 9.1e-10;
Matches 51; Conservative 24; Mismatches 49; Indels 21; Gaps 8;

| | | | |
|----|-----|--|-----|
| OY | 7 | PG-SSELSAFSAARTPLEGTSMAVTPDKYYVNIGDPDVATGQFCRCVPGARYPFSPT | 65 |
| | | : : : | : |
| DB | 110 | PGIPPELOIAFMASLATHSNONS-GIISSVETNIGNFDDVTGRFGAPVSGVFYFTPS | 168 |
| | | : : : | : |
| OY | 66 | AGKAPHKSLS---VMIVRNDEVQALAFDEQRPPGARPAASOSAMLQLDYGDWTWLR-H | 121 |
| | | : : : | : |
| DB | 169 | MMK--HEDVEEVYVLHMNGNTVSMSYEMK--GKSDTSSNHAAVLKLAKGDEVWLRMGN | 224 |
| | | : : : | : |
| OY | 122 | GAPHYALGAPG-----ATPFGYLIVY | 141 |
| | | : : : | : |
| DB | 225 | GALR-----GDHQRFSTFAGFLLP | 243 |
| | | : : : | : |

RESULT 15
C1QB_RAT STANDARD; PRT; 253 AA.
AC P31721;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Complement C1q subcomponent, B chain precursor.
GN C1QB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI TaxID=10116;

```

RN [1].
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Spleen;
RA MEDLINE=96062319; PubMed=7594503;
RX Schaeble W., Schaefer M.K.-H., Petry F., Fink T., Knebel D.,
RA Weihe E., Loos M.;
RT "Follicular dendritic cells, interdigitating cells, and cells of the
RT monocyte-macrophage lineage are the C1q-producing sources in the
RT spleen. Identification of specific cell types by in situ
RT hybridization and immunohistochemical analysis.";
RL J. Immunol. 155:4971-4978(1995).
RN [2].
RP SEQUENCE OF 71-79 AND 141-146.
RX MEDLINE=93218657; PubMed=8464426;
RA Wing M.G., Seilly D.J., Bridgman D.J., Harrison R.A.;
RT "Rapid isolation and biochemical characterization of rat C1 and C1q.";
RL Mol. Immunol. 30:433-440(1993).
CC -1- FUNCTION: C1Q ASSOCIATES WITH THE PROENZYMES C1R AND C1S TO YIELD
CC C1. THE FIRST COMPONENT OF THE SERUM COMPLEMENT SYSTEM. THE
CC COLLAGEN-LIKE REGIONS OF C1Q INTERACT WITH THE CA(2+)-DEPENDENT
CC C1R(2)C1S(2) PROENZYME COMPLEX, AND EFFICIENT ACTIVATION OF C1
CC TAKES PLACE ON INTERACTION OF THE GLOBULAR HEADS OF C1Q WITH THE
CC FC REGIONS OF IGG OR IGM ANTIBODY PRESENT IN IMMUNE COMPLEXES.
CC -1- SUBUNIT: C1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q, R
CC AND S IN THE MOLAR RATION OF 1:2:2. C1Q SUBCOMPONENT IS COMPOSED
CC OF NINE SUBUNITS, SIX OF WHICH ARE DISULFIDE-LINKED DIMERS OF THE
CC A AND B CHAINS, AND THREE OF WHICH ARE DISULFIDE-LINKED DIMERS OF
CC THE C CHAIN. IN ADDITION TO THE MAJOR A:B AND C:C DIMER BANDS,
CC RAT, UNLIKE HUMAN C1Q, CONTAINED MINOR DIMER SPECIES.
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C1q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X71127; CAA50440.1; -.
DR PIR; S49158; S49158.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
DR Complement pathway; Plasma; Hydroxylation; Glycoprotein; Collagen;
KW Repeat; Signal.
KW
FT SIGNAL 1 25 BY SIMILARITY.
FT CHAIN 26 253 COMPLEMENT C1Q SUBCOMPONENT, B CHAIN.
FT DOMAIN 29 112 COLLAGEN-LIKE..
FT DOMAIN 113 253 C1Q.
FT DISULFID 29 29 INTERCHAIN (WITH C-26 IN CHAIN A).
FT MOD_RES 33 33 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 36 36 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 39 39 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 51 51 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 54 54 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 57 57 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 60 60 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 63 63 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 75 75 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 81 81 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 84 84 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 90 90 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 96 96 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 99 99 HYDROXYLATION (BY SIMILARITY).
SQ SEQUENCE 253 AA; 26589 MW; 1CB40622571BFC9B CRC64;

```


Best Local Similarity 29.2%; Pred. No. 9.4e-10;
Matches 45; Conservative 36; Mismatches 58; Indels 15; Gaps 7;

```
QY      2 GPTPGP-----GSSELR-----AFSARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQF 52
      |||||  ||:  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db     100 GP-PCPRGPKGSGDYKATQKVAFSALRTVNSALRPNOAIRFEKVITNVNDNYPEPRSGKF 158
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      53 RCRVPGAYFFSFTAGKAPHKSLVNLVRNDEVQ-ALAFDEQRPGARRAASQSAMLQID 111
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db     159 TCKVPGILYYFTYHASSRGNLCVNIVRGRDRDMQKVLTFCDYAO-NTFQVTTGGVVLKLE 217
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      112 YGDTVWLRLHGAPHYALGAPGAT--ESGYLVYAD 143
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db     218 QEEVVHLQATD-KNSLLGVEGANSIFTGFLLFPD 250
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```

Search completed: January 12, 2004, 08:12:58
Job time : 6.13772 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:10 ; Search time 19.6946 Seconds
(without alignments)
1873.686 Million cell updates/sec

Title: US-10-085-167-2_COPY_17_159
Perfect score: 745
Sequence: 1 LGPTPGGSSELSRASFSAAR.....PHYALGAPGATFSGYLVYAD 143

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 737 | 98.9 | 329 | 4 Q8IV25 | Q8IV25 homo sapien |
| 2 | 706.5 | 94.8 | 326 | 11 Q8R066 | Q8R066 mus musculu |
| 3 | 355 | 47.7 | 205 | 11 Q9D0W2 | Q9D0W2 mus musculu |
| 4 | 349 | 46.8 | 205 | 11 Q9DCB6 | Q9DCB6 mus musculu |
| 5 | 222 | 29.8 | 158 | 4 Q9H667 | Q9H667 homo sapien |
| 6 | 222 | 29.8 | 1077 | 4 Q8TE71 | Q8TE71 homo sapien |
| 7 | 216 | 29.0 | 158 | 11 Q8K110 | Q8K110 mus musculu |
| 8 | 200 | 26.8 | 294 | 11 Q9D8U4 | Q9D8U4 mus musculu |
| 9 | 192 | 25.8 | 243 | 6 Q95JD7 | Q95JD7 macaca mula |
| 10 | 189 | 25.4 | 243 | 4 Q8N6P2 | Q8N6P2 homo sapien |
| 11 | 185 | 24.8 | 247 | 11 Q8BRW2 | Q8BRW2 mus musculu |
| 12 | 181.5 | 24.4 | 243 | 11 Q8R002 | Q8R002 mus musculu |
| 13 | 180.5 | 24.2 | 264 | 11 Q8BKRO | Q8BKRO mus musculu |
| 14 | 180 | 24.2 | 244 | 11 Q8K3R4 | Q8K3R4 rattus norv |
| 15 | 179 | 24.0 | 240 | 6 Q95M04 | Q95M04 bos taurus |
| 16 | 178.5 | 24.0 | 243 | 11 Q8K479 | Q8K479 mus musculu |

| | | | | | |
|----|-------|------|------|-----------|---------------------|
| 17 | 171.5 | 23.0 | 312 | 11 Q8CHX9 | Q8CHX9 mus musculu |
| 18 | 171 | 23.0 | 287 | 11 Q8CFR0 | Q8CFR0 mus musculu |
| 19 | 165.5 | 22.2 | 194 | 6 Q95J95 | Q95J95 canis famil |
| 20 | 164.5 | 22.1 | 120 | 11 Q8R1Z2 | Q8R1Z2 mus musculu |
| 21 | 162.5 | 21.8 | 246 | 11 Q9ES30 | Q9ES30 mus musculu |
| 22 | 162.5 | 21.8 | 289 | 11 Q8BVD7 | Q8BVD7 mus musculu |
| 23 | 160.5 | 21.5 | 182 | 11 Q8R1P2 | Q8R1P2 mus musculu |
| 24 | 160.5 | 21.5 | 281 | 11 Q9QXP7 | Q9QXP7 mus musculu |
| 25 | 156 | 20.9 | 245 | 11 Q9DCM6 | Q9DCM6 mus musculu |
| 26 | 150 | 20.1 | 224 | 4 Q8IUK8 | Q8IUK8 homo sapien |
| 27 | 150 | 20.1 | 224 | 11 Q8BGU2 | Q8BGU2 mus musculu |
| 28 | 146.5 | 19.7 | 197 | 11 Q9JHG0 | Q9JHG0 mus musculu |
| 29 | 144.5 | 19.4 | 333 | 4 Q8IUV4 | Q8IUV4 homo sapien |
| 30 | 142.5 | 19.1 | 173 | 6 Q62789 | Q62789 sus scrofa |
| 31 | 141 | 18.9 | 744 | 11 Q921S8 | Q921S8 mus musculu |
| 32 | 141 | 18.9 | 744 | 11 Q8BGL6 | Q8BGL6 mus musculu |
| 33 | 139 | 18.7 | 675 | 6 Q9N178 | Q9N178 sus scrofa |
| 34 | 135.5 | 18.2 | 705 | 4 Q8TEJ5 | Q8TEJ5 homo sapien |
| 35 | 134.5 | 18.1 | 213 | 5 P83425 | P83425 mytilus edu |
| 36 | 132.5 | 17.8 | 195 | 11 Q8BZS3 | Q8BZS3 mus musculu |
| 37 | 131 | 17.6 | 185 | 5 Q9GQV4 | Q9GQV4 stronglyloce |
| 38 | 130 | 17.4 | 222 | 13 Q8J126 | Q8J126 salvelinus |
| 39 | 127.5 | 17.1 | 102 | 6 Q95MB2 | Q95MB2 equus cabal |
| 40 | 127.5 | 17.1 | 295 | 11 Q9Z1K4 | Q9Z1K4 rattus norv |
| 41 | 127.5 | 17.1 | 1017 | 11 Q99K41 | Q99K41 mus musculu |
| 42 | 127 | 17.0 | 198 | 11 Q8BMF0 | Q8BMF0 mus musculu |
| 43 | 126 | 16.9 | 198 | 11 Q8BME9 | Q8BME9 mus musculu |
| 44 | 121.5 | 16.3 | 347 | 4 Q96IH6 | Q96IH6 homo sapien |
| 45 | 121.5 | 16.3 | 583 | 4 Q96G58 | Q96G58 homo sapien |

ALIGNMENTS

RESULT 1
Q8IV25 PRELIMINARY; PRT; 329 AA.
ID Q8IV25
AC Q8IV25;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Similar to C1q and tumor necrosis factor related protein 4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC035628; AAH35628.1; -
SQ SEQUENCE 329 AA; 35256 MW; 16064DA8182A6732 CRC64;

Query Match 98.9%; Score 737; DB 4; Length 329;
Best Local Similarity 99.3%; Pred. No. 1.3e-68;
Matches 142; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

| | |
|----------|--|
| Qy | 1 LGPTPGGSSELSRASFSAARTTPLEGSEMAVTFDKVYVNIIGDFDVATGQFRCRVPAY 60 |
| Db | 17 LGPTPGGSSELSRASFSAARTTPLEGSEMAVTFDKVYVNIIGDFDVATGQFRCRVPAY 76 |
| Qy | 61 FFSTAGKAPHKSLVMLVRNDEVALAFDEQRRPGARRAASQASAMQLDYGDTVWLRL 120 |
| Db | 77 FFSTAGKAPHKSLVMLVRNDEVALAFDEQRRPGARRAASQASAMQLDYGDTVWLRL 136 |
| Qy | 121 HGAPHYALGAPGATFSGYLVYAD 143 |
| Db | 137 HGAPHYALGAPGATFSGYLVYAD 159 |
| RESULT 2 | |
| Q8R066 | |

| ID | Q8R066 | PRELIMINARY; | PRT; | 326 AA. |
|----------|--|--------------|------------------|--------------------|
| AC | Q8R066; | | | |
| DT | 01-JUN-2002 (TREMBLrel. 21, Created) | | | |
| DT | 01-JUN-2002 (TREMBLrel. 21, Last sequence update) | | | |
| DT | 01-MAR-2003 (TREMBLrel. 23, Last annotation update) | | | |
| DE | Similar to Clq and tumor necrosis factor related protein 4. 0710001E10RIK. | | | |
| GN | Mus musculus (Mouse). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | |
| OX | NCBI_TaxID=10090; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | Strauberg R.; | | | |
| RL | Submitted (APR-2002) to the EMBL/GenBank/DBJ databases. | | | |
| DR | EMBL; BC027315; AAH27315.1; - | | | |
| DR | MGI; MGI:1914695; 0710001E10RIK. | | | |
| DR | InterPro; IPR001073; Clq. | | | |
| DR | Pfam; PF00386; Clq; 2. | | | |
| DR | SMART; SM00110; Clq; 2. | | | |
| DR | PROSITE; PS01113; Clq; 2. | | | |
| DR | SEQUENCE 326 AA; 35057 MW; 72339172B7B1051A CRC64; | | | |
| QY | Query Match | 94.8%; | Score 706.5; | DB 11; Length 326; |
| | Best Local Similarity | 96.5%; | Pred. No. 2e-65; | |
| | Matches 138; Conservative 1; Mismatches 3; Indels 1; Gaps 1; | | | |
| Db | 1 LGPTPPGSSSELSRASFSAARTTPLEGTSEMAVTFDKVYVNIIGDPVATGQFRCRVPAY 60 | | | |
| | 17 LGPA-GPSSSELSRASFSAARTTPLEGTSEMAVTFDKVYVNIIGDPVATGQFRCRVPAY 75 | | | |
| QY | 61 FFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTVWLRL 120 | | | |
| | 76 FFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTVWLRL 135 | | | |
| Db | 121 HGAPHYALGAPGATFSGYLVYAD 143 | | | |
| | 136 HGAPQYALGAPGATFSGYLVYAD 158 | | | |
| QY | 121 HGAPHYALGAPGATFSGYLVYAD 143 | | | |
| Db | 136 HGAPQYALGAPGATFSGYLVYAD 158 | | | |
| RESULT 3 | | | | |
| Q9D0W2 | | | | |
| ID | Q9D0W2 | PRELIMINARY; | PRT; | 205 AA. |
| AC | Q9D0W2; | | | |
| DT | 01-JUN-2001 (TREMBLrel. 17, Created) | | | |
| DT | 01-JUN-2001 (TREMBLrel. 17, Last sequence update) | | | |
| DT | 01-OCT-2002 (TREMBLrel. 22, Last annotation update) | | | |
| DE | 0710001E10RIK protein. | | | |
| GN | 0710001E10RIK. | | | |
| OS | Mus musculus (Mouse). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | |
| OX | NCBI_TaxID=10090; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | STRAIN=C57BL/6J; TISSUE=Embryo; | | | |
| RC | MEDLINE=21085660; PubMed=11217851; | | | |
| RA | Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., | | | |
| RA | Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., | | | |
| RA | Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I., | | | |
| RA | Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., | | | |
| RA | Kadota K., Matsuda H. A., Ashburner M., Batalov S., Casavant T., | | | |
| RA | Fleischmann W., Gaasterland T., Gliszi C., King B., Kochiwa H., | | | |
| RA | Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J., | | | |
| RA | Schriml L. M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., | | | |
| RA | Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., | | | |
| RA | Blake J., Buffelli D., Bojunga N., Carninci P., de Bonaldo M. F., | | | |
| RA | Brownstein M. J., Bult C., Fletcher C., Fujita M., Gariboldi M., | | | |
| RA | Guestincich S., Hill D., Hofmann M., Hume D. A., Kamiya M., Lee N. H., | | | |
| RA | Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., | | | |
| RA | Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., | | | |
| RA | Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., | | | |
| RA | Suzuki H., Toyo-oka K., Wang K. H., Weitz C., Wittaker C., Wilming L., | | | |

```
RA W ynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohatsu S.,  
RA Hayashizaki Y.;  
RT "functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
DR EMBL; AK004340; BAB23268.1; -;  
DR MGD; MGI:1914695; 0710001E1ORik.  
DR InterPro; IPR001073; Clq.  
DR Pfam; PF00386; Clq; 1.  
DR PRINTS; PR00007; COMPLEMENTClQ.  
DR SMART; SM00110; ClQ; 1.  
DR PROSITE; PS01113; ClQ; 1.  
SQ SEQUENCE 205 AA; 22190 MW; B9B237793C54786D CRC64;
```

Query Match 47.7%; Score 355; DB 11; Length 205;
Best Local Similarity 53.3%; Pred. No. 5.3e-29;
Matches 81; Conservative 12; Mismatches 45; Indels 14; Gaps 4.

| | | | | |
|----|--|-----|---|-----|
| OY | | 3 | PTPGSSSELRSAFSAARTTPLEGTS-----EMAVTFDKVYNIGGEFDVATGOFRRCRV | 56 |
| | | : | : : | : |
| Dd | | 43 | PARGPAPEPRSASFSAARSLVGSDAAGPHRRPLAFDTBLVNIGDFDAAGVFRCRL | 102 |
| OY | | 57 | PGAYFFSF TAGKAPHKSLSVMLVRNRDEVQALAFDE--QRPGAARRAASOSAMLQLDYGD | 114 |
| | | : | : : | : |
| Dd | | 103 | PGAYFFSFTLGKLPRKTLSVKLMKNRDEVQAITYDDGASRR---REMGSQSVM LPLRRGD | 159 |
| OY | | 115 | TWVLRLH--GAPHYALGADGATFSGLYLVAD | 143 |
| | | : | : | : |
| Dd | | 160 | AWWLSHDHG DGAYSNGKYITFSGLVPDP | 191 |

```

RESULT 4
Q9DCB6 PRELIMINARY; PRT; 205 AA.
ID Q9DCB6;
AC Q9DCB6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE 0710001E10Rik protein.
GN 0710001E10Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Caaveant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Wittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK002948; BAB22473.1; -
DR MGD; MGI:1914695; 0710001E10Rik.
DR InterPro; IPR001073; Clq.
DR Pfam; PF00386; Clq; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.

```

DR PROSITE; PS01113; ClQ; 1.
SQ SEQUENCE 205 AA; 22215 MW; 56AD37793C437300 CRC64;
Query Match
Best Local Similarity 46.8%; Score 349; DB 11; Length 205;
Matches 80; Conservative 12; Mismatches 46; Indels 14; Gaps 4;
QY 3 PTPGSGSELRSAFSAARTTPL-EGTSEMAVTFDKVYVNIIGDFDVATGQFRGCRV 56
DB 43 PARGPAAPEPRSAFSAARTSRIVGSDAAGPRHRPLAFDTLVLNIGDFDAAGVFRCL 102
QY 57 PGAYFFSFTAGKAPHKSLSVMLVRNDEVQALAFDE--QRRPGARRAASQASAMLQLDYGD 114
DB 103 PGAYFFSFTLGLKLPKRTLSTVKLMKNRDEVQAMITYDDGASRR---REMOSQSVRLPLRGD 159
QY 115 TVWLRLH---GAPHYALGAPGATFSGYLVYAD 143
DB 160 AVWLISHDHDGYGAYSNHGKTYTFSGFLVYPD 191

RESULT 5

Q9H667 PRELIMINARY; PRT; 158 AA.
ID Q9H667
AC Q9H667;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ22569.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK026222; BAB15398.1; -
DR EMBL; BC007520; AAH07520.1; -
DR InterPro: IPR001073; ClQ.
DR Pfam; PF00386; ClQ; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; ClQ; 1.
DR PROSITE; PS01113; ClQ; 1.
KW Hypothetical protein.
SQ SEQUENCE 158 AA; 17625 MW; 47DB10EDD6DC9760 CRC64;

Query Match 29.8%; Score 222; DB 4; Length 158;
Best Local Similarity 41.7%; Pred. No. 2.9e-15;
Matches 58; Conservative 18; Mismatches 59; Indels 4; Gaps 4;
QY 7 PGSSELRSAFSAARTTPL-EGTSEMAVTFDKVYVNIIGDFDVATGQFRGCRVPGAYFFSFT 65
DB 22 PLPQOMRVAFSAARTSNLAPGTLDPPIVFDLLNLNGETFDLQGRFNCVPVNGTYVFIFH 81
QY 66 AGK-APHKLSVMLVRNDEVQALAFDEQRRPGARRAASQASAMLQLDYGDYTWLRLHGAP 124
DB 82 MLKLAVNVPLVYVNLKKN-EEVLVSAVYANDGAPD-HETASNHAILQLFGDQIWLRLHRGA 139
QY 125 HYALGAPGATFSGYLVYAD 143
DB 140 IYGSSWKYSTFSGYLLYQD 158

RESULT 6

O8TE71 PRELIMINARY; PRT; 1077 AA.
ID O8TE71
AC O8TE71;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE BEG1L.
GN BEG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Aerbajnal W., Miller J.L.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY074490; AAL71549.1; -
DR InterPro: IPR001073; ClQ.
DR Pfam; PF00386; ClQ; 1.
DR SMART; SM00110; ClQ; 1.
DR PROSITE; PS01113; ClQ; 1.
SQ SEQUENCE 1077 AA; 120974 MW; 2B88BF3C47D032D6 CRC64;

Query Match 29.8%; Score 222; DB 4; Length 1077;
Best Local Similarity 41.7%; Pred. No. 3.4e-14;
Matches 58; Conservative 18; Mismatches 59; Indels 4; Gaps 4;

QY 7 PGSSELRSAFSAARTTPL-EGTSEMAVTFDKVYVNIIGDFDVATGQFRGCRVPGAYFFSFT 65
DB 941 PLPQOMRVAFSAARTSNLAPGTLDPPIVFDLLNLNGETFDLQGRFNCVPVNGTYVFIFH 1000
QY 66 AGK-APHKLSVMLVRNDEVQALAFDEQRRPGARRAASQASAMLQLDYGDYTWLRLHGAP 124
DB 1001 MLKLAVNVPLVYVNLKKN-EEVLVSAVYANDGAPD-HETASNHAILQLFGDQIWLRLHRGA 1058
QY 125 HYALGAPGATFSGYLVYAD 143
DB 1059 IYGSSWKYSTFSGYLLYQD 1077

RESULT 7

O8K110 PRELIMINARY; PRT; 158 AA.
ID O8K110
AC O8K110;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Similar to hypothetical protein FLJ22569.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027523; AAH27523.1; -
DR InterPro: IPR001073; ClQ.
DR Pfam; PF00386; ClQ; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; ClQ; 1.
DR PROSITE; PS01113; ClQ; 1.
KW Hypothetical protein.
SQ SEQUENCE 158 AA; 17533 MW; 86E9321C99225FCB CRC64;

Query Match 29.0%; Score 216; DB 11; Length 158;
Best Local Similarity 41.0%; Pred. No. 1.2e-14;
Matches 57; Conservative 19; Mismatches 59; Indels 4; Gaps 4;

QY 7 PGSSELRSAFSAARTTPL-EGTSEMAVTFDKVYVNIIGDFDVATGQFRGCRVPGAYFFSFT 65
DB 22 PLPQOMRVAFSAARTSNLAPGTLDPPIVFDLLNLNGETFDLQGRFNCVPVNGTYVFIFH 81

QY 66 AGK-APHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDYDGTWLRHGAP 124
Db 82 MLKLAIVNPLVYVLMKN-EENVVSAVYANDGAPD-HETASNHAVLQLQGDQIMLRHGA 139
QY 125 HVALGAPGATFSGYLVYAD 143
Db 140 IYSSMWKYSTFSGYLVYQD 158

RESULT 8

Q9D8U4 PRELIMINARY; PRT; 294 AA.
AC Q9D8U4;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE 1810033K05R1k protein (RIKEN CDNA 1810033K05 gene).
GN 1810033K05R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staedtli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyoko-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.

RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK007683; BAB25187.1; -
DR EMBL; BC030324; AAH30324.1; -
DR MGD; MGI:1916433; 1810033K05R1k.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 2.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR PROSITE; PS01113; C1Q; 1.
SQ SEQUENCE 294 AA; 30865 MW; 6D3905AE7C19E6FA CRC64;

Query Match 26.8%; Score 200; DB 11; Length 294;
Best Local Similarity 34.9%; Pred. No. 1.3e-12;
Matches 53; Conservative 30; Mismatches 45; Indels 24; Gaps 8;

QY 5 GGP--GSSELRSAPSA--RTPLEGTSEMAVTFDKVYVNIIGDFVATGQRCRVPGA 59
Db 147 GPCSCGSSRAKSAFSAVATKSYPRE--RUPIKFDKILMNEGHHYASSGKFCVSPGI 203
QY 60 YPFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDYDGTWLR 119
Db 204 YFTYDITLA-NKHLAIGLVHN-GQYRIRTPD--ANTGNHVDVASSGSTITALKEGDEVWLQ 259

QY 120 LHGA-----PHYALGAPGATFSGYLVYAD 143
Db 260 IFYSEQNGLFYDPYWT----DSLFTGFLIYAD 287

RESULT 9

Q95JD7 PRELIMINARY; PRT; 243 AA.
AC Q95JD7;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Adiponectin.
GN APM1.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.

RC TISSUE=Adipose tissue;
RX MEDLINE=21232234; PubMed=11334417;
RA Hotta K., Funahashi T., Bodkin N.L., Ortmeier H.K., Arita Y.,
RA Hansen B.C., Matsuzawa Y.;
RT "Circulating concentrations of the adipocyte protein adiponectin are
RT decreased in parallel with reduced insulin sensitivity during the
RT progression to type 2 diabetes in rhesus monkeys."
RL Diabetes 50:1126-1133(2001).
DR EMBL; AF404407; AAK92202.1; -
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR ProDom; PD000007; Collagen; 1.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Collagen.
SQ SEQUENCE 243 AA; 26264 MW; 49A45DAF2B4613FD CRC64;

Query Match 25.8%; Score 192; DB 6; Length 243;
Best Local Similarity 31.8%; Pred. No. 6.8e-12;
Matches 47; Conservative 29; Mismatches 54; Indels 18; Gaps 6;

QY 5 GPGSSLSRASFSAARTPLEGTSEMAVTFDKVYVNIIGDFVATGQRCRVPGAYFSF 64
Db 103 PGECAVYVRSASFVGLFTYVT-VPNMPIRFTKIFVQONHYDVGSTGKFNIPGLYFYAY 161
QY 65 TAGKAPH-----KSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDYDGTWLR 119
Db 162 -----HITVYMKDVKVSLEK-KDKAMLFTYDQYQENNVQD-ASGSVLLHLEVGQVWLQ 213
QY 120 LHGAH----YALGAPGATFSGYLVYAD 143
Db 214 VYGEGERNGLYADNDNDSTFTGFLLYHD 241

RESULT 10

Q9N6P2 PRELIMINARY; PRT; 243 AA.

AC Q9N6P2;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE C1q and tumor necrosis factor related protein 5.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

Matches 56; Conservative 15; Mismatches 61; Indels 25; Gaps 6;
QY 2 GPTPGPGSS--ELRSAFSAARTTPLEGTSE-MAVTFDKVYVNIIGDDVATGQRCRVPG 58
Db 116 GQAGSPGSSCQTHYSAFSVGRKTGLHSENFSLFDRVFVNTDGHFDMATGSAFVAPLRG 175
QY 59 AYFFSFTAGKAPHKSLSLVLRNRDEVOALAFDEQRRPGARR-AAQSQAMLQLDYGDVTW 117
Db 176 LYFFSLNVHSMYKETYVHIVHNEQAVVIL---YAQPSERSIMQSQSVMLPLVPGRVW 231
QY 118 LRL-----HGAPHYALGAPGATFSGYLVYAD 143
Db 232 VRLFKRERENGISDDVDVTY-----ITFSGHLIKAE 262

RESULT 14
Q8K3R4 PRELIMINARY; PRT; 244 AA.
AC Q8K3R4;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)
DE 30 kDa adipocyte complement-related protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Berg A.H., Scherer P.E.;
RT "Cloning of the rat homolog to murine adipocyte complement-related
RT protein of 30 kDa."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY033885; AAK61608.1; -.
DR InterPro; IPR001073; Clq.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; Clq; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR ProDom; PD000007; Collagen; 1.
DR SMART; SM00110; Clq; 1.
DR PROSITE; PS01113; Clq; 1.
KW Collagen.
SQ SEQUENCE 244 AA; 26410 MW; 75B2B1DF68E2633E CRC64;

Query Match 24.2%; Score 180; DB 11; Length 244;
Best Local Similarity 30.4%; Pred. No. 1.2e-10;
Matches 45; Conservative 29; Mismatches 56; Indels 18; Gaps 6;

QY 5 PGPSSSELRSAFSAARTTPLEGTSEMAVTFDKVYVNIIGDDVATGQRCRVPGAYFFSF 64
Db 104 PGEAAYMYHSAFSGLETRVT-VPNVPIRFTKIFYNQNHYDSTGKFCNIPGLYYFSY 162
QY 65 TAGKAPH-----KSLSVMLVRNRDEVOALAFDEQRRPGARRAASQSAMQLDYGDVTWLR 119
Db 163 -----HITVYMKDVKVSLEK-KDKAVLFTYDQYQEKVNDQ-ASGSMLHLLEVGDQVWLQ 214
QY 120 LHGAPH---YALGAPGATFSGYLVYAD 143
Db 215 VYEGEDNNGLYADNVNDSTFTGFLLYHD 242

RESULT 15
Q95MQ4 PRELIMINARY; PRT; 240 AA.
AC Q95MQ4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)
DE Adipose tissue-specific protein adipo Q.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21369933; Pubmed=11382781;
RA Sato C., Yasukawa Z., Honda N., Matsuda T., Kitajima K.;
RT "Identification and Adipocyte Differentiation-dependent Expression of
RT the Unique Disialic Acid Residue in an Adipose Tissue-specific
RT Glycoprotein, Adipo Q."
RL J. Biol. Chem. 276:28849-28856(2001).
DR EMBL; AF269230; AAK58902.1; -.
DR InterPro; IPR001073; Clq.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; Clq; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR ProDom; PD000007; Collagen; 1.
DR SMART; SM00110; Clq; 1.
DR PROSITE; PS01113; Clq; 1.
KW Collagen.
SQ SEQUENCE 240 AA; 26091 MW; C6253BA803B9A668 CRC64;

Query Match 24.0%; Score 179; DB 6; Length 240;
Best Local Similarity 31.5%; Pred. No. 1.5e-10;
Matches 46; Conservative 29; Mismatches 53; Indels 18; Gaps 7;

QY 5 PGPSSSELRSAFSAARTTPLEGTSEMAVTFDKVYVNIIGDDVATGQRCRVPGAYFFSF 64
Db 99 PGEAAYVYRSAFSGLETRVT-VPNVPIRFTKIFYNQNHYDSTGKFCNIPGLYYFSY 157
QY 65 TAGKAPH-----KSLSVMLVRNRDEVOALAFDEQRRPGARRAASQSAMQLDYGDVTWLR 119
Db 158 -----HITVYMKDVKVSLEK-KDKAVLFTYDQYQEKVNDQ-ASGSVLLHLEVGDQVWLQ 209
QY 120 LH-GAPH---YALGAPGATFSGYLVY 141
Db 210 VYEGENHNGVYADNVNDSTFTGFLLY 235

Search completed: January 12, 2004, 08:19:09
Job time : 20.6946 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:09 ; Search time 21.5499 Seconds
(without alignments)
1053.272 Million cell updates/sec

Title: US-10-085-167-2_COPY_17_159
Perfect score: 745
Sequence: 1 LGPTPGSSSELRSFAFSAR.....PHYALGAPGATFSGYLVIAD 143

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|--------------------|
| 1 | 745 | 100.0 | 299 | 23 | ABG70385 | Adipocyte compleme |
| 2 | 745 | 100.0 | 329 | 22 | AAB61606 | Human ZACRP4. Hom |
| 3 | 737 | 98.9 | 284 | 23 | ABG70384 | Adipocyte compleme |
| 4 | 737 | 98.9 | 348 | 22 | AAB61423 | Human MANGO 245 pr |
| 5 | 733 | 98.4 | 334 | 22 | AAB61466 | Human MANGO 245 ma |
| 6 | 726 | 97.4 | 329 | 23 | ABG79643 | Human novel secret |
| 7 | 708 | 95.0 | 329 | 22 | AAB61424 | Monkey MANGO 245 p |
| 8 | 707.5 | 95.0 | 199 | 22 | AAB61488 | Murine MANGO 245 p |
| 9 | 646 | 86.7 | 126 | 22 | AAB61469 | Human MANGO 245 C1 |

| | | | | | | |
|----|-------|------|------|----|----------|--------------------|
| 10 | 639 | 85.8 | 126 | 22 | AAB61489 | Murine MANGO 245 C |
| 11 | 635 | 85.2 | 125 | 22 | AAB61477 | Human MANGO 245 ex |
| 12 | 622 | 83.5 | 126 | 22 | AAB61472 | Monkey MANGO 245 C |
| 13 | 618 | 83.0 | 123 | 21 | AAB42189 | Human OREF ORF1953 |
| 14 | 448 | 60.1 | 221 | 23 | ABG70383 | Adipocyte compleme |
| 15 | 426.5 | 57.2 | 130 | 22 | AAB61468 | Mature monkey MANG |
| 16 | 341.5 | 45.8 | 133 | 23 | ABG70386 | Adipocyte compleme |
| 17 | 320 | 43.0 | 192 | 22 | AAB61479 | Human MANGO 245 cy |
| 18 | 314 | 42.1 | 134 | 22 | AAB61473 | Monkey MANGO 245 C |
| 19 | 305 | 40.9 | 127 | 22 | ABB11578 | Human secreted pro |
| 20 | 301.5 | 40.5 | 117 | 22 | AAB61470 | Human MANGO 245 C1 |
| 21 | 222 | 29.8 | 202 | 22 | AAM99927 | Human polypeptide |
| 22 | 222 | 29.8 | 709 | 23 | ABB08645 | Human pancreatic c |
| 23 | 222 | 29.8 | 710 | 23 | AAO15420 | Human genset metab |
| 24 | 222 | 29.8 | 746 | 20 | AAV29512 | Human lung tumour |
| 25 | 222 | 29.8 | 746 | 21 | AAB44461 | Human lung tumour |
| 26 | 222 | 29.8 | 746 | 22 | AAE13802 | Human lung tumour |
| 27 | 222 | 29.8 | 800 | 22 | AAV78334 | Human lung tumour |
| 28 | 222 | 29.8 | 908 | 22 | AAE09839 | Human protein SEQ |
| 29 | 222 | 29.8 | 909 | 22 | AAE09845 | Novel human protei |
| 30 | 222 | 29.8 | 957 | 22 | AAE09841 | Novel human protei |
| 31 | 222 | 29.8 | 958 | 22 | AAE09847 | Novel human protei |
| 32 | 222 | 29.8 | 992 | 22 | AAE09843 | Novel human protei |
| 33 | 222 | 29.8 | 993 | 22 | AAE09849 | Novel human protei |
| 34 | 222 | 29.8 | 1042 | 22 | AAE09838 | Novel human protei |
| 35 | 222 | 29.8 | 1043 | 22 | AAE09844 | Novel human protei |
| 36 | 222 | 29.8 | 1091 | 22 | AAE09840 | Novel human protei |
| 37 | 222 | 29.8 | 1092 | 22 | AAE09846 | Novel human protei |
| 38 | 222 | 29.8 | 1126 | 22 | AAE09842 | Novel human protei |
| 39 | 222 | 29.8 | 1127 | 22 | AAE09848 | Novel human protei |
| 40 | 209 | 28.1 | 710 | 22 | AAV78335 | Human protein SEQ |
| 41 | 200 | 26.8 | 285 | 21 | AAV37344 | Murine ACRP30R1M. |
| 42 | 200 | 26.8 | 294 | 21 | AAV76039 | Rat skin cell prot |
| 43 | 200 | 26.8 | 294 | 22 | AAE55978 | Skin cell protein |
| 44 | 200 | 26.8 | 294 | 23 | ABE72178 | Rat protein isolat |
| 45 | 197 | 26.4 | 285 | 21 | AAV76103 | Rat skin cell secr |

ALIGNMENTS

| | |
|----------|---|
| RESULT 1 | |
| ABG70385 | |
| ID | ABG70385 standard; Protein; 299 AA. |
| XX | |
| AC | ABG70385; |
| XX | |
| DT | 05-NOV-2002 (first entry) |
| XX | |
| DE | Adipocyte complement-related C1q Tumour Necrosis Factor-like protein #3. |
| XX | |
| KW | Human; NOX; NOX-associated disorder; cardiomyopathy; atherosclerosis; |
| KW | cell signal processing; metabolic pathway modulation; metabolic disorder; |
| KW | obesity; diabetes; infectious disease; neurodegenerative disorder; acne; |
| KW | Alzheimer's disease; Parkinson's disease; immune disorder; cancer; |
| KW | haematopoietic disorder; cirrhosis; pancreatitis; learning defect; |
| KW | memory defect; infertility; congenital heart defect; hair growth; |
| KW | pigmentation disorder; endocrine disorder; respiratory disease; health; |
| KW | gastro-intestinal disease; reproductive; neurological disease; |
| KW | bone marrow transplantation; endocrine disease; allergy; inflammation; |
| KW | neurological disorder; urinary system disorder; age-related disorder; |
| KW | neuropsychiatric disorder; EGF-related protein; SCUB1; TEN-M4; |
| KW | adipocyte complement-related C1q tumour necrosis factor; out at first; |
| KW | beta adrenergic receptor kinase; EphA6/enk-2; glucose transporter; |
| KW | type Ia membrane sushi-containing domain; butyrophillin; |
| KW | type Ia membrane-sushi domain containing. |
| OS | Homo sapiens. |
| XX | |
| XX | |
| PN | WO200257453-A2; |
| XX | |
| PD | 25-JUL-2002. |
| XX | |

PF 19-DEC-2001; 2001WO-US50331.
XX
PR 19-DEC-2000; 2000US-265704P.
PR 20-DEC-2000; 2000US-257314P.
PR 02-MAY-2001; 2001US-288153P.
PR 29-MAY-2001; 2001US-294075P.
PR 24-JUL-2001; 2001US-307506P.
PR 10-AUG-2001; 2001US-311590P.
PR 10-AUG-2001; 2001US-311613P.
PR 29-AUG-2001; 2001US-315617P.
PR 14-SEP-2001; 2001US-322358P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Gangoli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
PI Stone DJ, Anderson D, Shinkets RA, Burgess CE, Zerhusen BD, Liu X;
PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
DR WPI; 2002-590744/63.
DR N-PSDB; ABS52097.
XX
PT Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious
PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
PT cancer -
XX
PS Claim 1; Page 31; 318pp; English.
XX
CC The present invention relates to new NOVX polypeptides. The invention is
CC useful for treating or preventing a NOVX-associated disorder such as
CC cardiomyopathy or atherosclerosis, where the disorder is related to cell
CC signal processing and metabolic pathway modulation in a subject,
CC preferably human. The invention is also useful for treating metabolic
CC disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
CC disorders, haematopoietic disorders and various cancers. The molecules of
CC the invention are also useful for treating or preventing cirrhosis,
CC pancreatitis, learning and memory defects, infertility, congenital heart
CC defects, acne, hair growth, pigmentation disorders, endocrine disorders,
CC respiratory disease, gastro-intestinal diseases, reproductive, health,
CC neurological diseases, bone marrow transplantation, endocrine diseases,
CC allergy and inflammation, nephrological disorders, urinary system
CC disorders, neuropsychiatric disorders and age-related disorders.
CC The present amino acid sequence represents a NOVX protein of the
CC invention.
XX
SQ Sequence 299 AA;
Query Match 100.0%; Score 745; DB 23; Length 299;
Best Local Similarity 100.0%; Pred. No. 1.5e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LGPTPGSGSELRSAFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPAY 60
Db 17 LGPTPGSGSELRSAFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPAY 76
QY 61 FFSTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTWRL 120
Db 77 FFSTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTWRL 136
QY 121 HGAPHYALGAPGATFSGYLVYAD 143
Db 137 HGAPHYALGAPGATFSGYLVYAD 159

RESULT 2
AAB61606
ID AAB61606 standard; Protein; 329 AA.
XX
AC AAB61606;
XX
DT 05-APR-2001 (first entry)
XX

DE Human ZACRP4.
XX
KW Human; zacrpa; complement factor C1q domain; chromosome 11q11;
KW energy balance; cellular metabolic reaction; autocrine factor;
KW development; cell proliferation; differentiation; cell survival.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..16
FT Domain /label= Signal_peptide
FT Domain /label= C1q_domain_#1
FT Domain /label= C1q_domain_#2
FT Domain /label= C1q_domain_#2
XX
PN WO200102565-A2.
XX
PD 11-JAN-2001.
XX
PE 28-JUN-2000; 2000WO-US17692.
XX
PR 01-JUL-1999; 99US-0346502.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Holloway JL, Lok S;
XX
DR WPI; 2001-138140/14.
DR N-PSDB; AAF28672.
XX
PT Novel secreted protein ZACRP4 polypeptides having tandem C1q globular
PT domains, useful for studying cell-cell communication and regulation of
PT cellular processes -
XX
PS Claim 1; Page 77-78; 82pp; English.
XX
CC The present sequence is human ZACRP4 protein. ZACRP4 protein has two
CC complement factor C1q domains. The ZACRP4 gene is located on human
CC chromosome 11q11. The ZACRP4 coding sequence and protein have a number of
CC uses described in the specification, including, modulation of energy
CC balance and cellular metabolic reactions in mammals. In addition, ZACRP4
CC protein is useful as an autocrine factor, particularly during
CC development, in mediating the processes of an organism, in regulating
CC cellular processes such as cell proliferation and/or differentiation,
CC cell survival and energy balance.
XX
SQ Sequence 329 AA;
Query Match 100.0%; Score 745; DB 22; Length 329;
Best Local Similarity 100.0%; Pred. No. 1.7e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LGPTPGSGSELRSAFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPAY 60
Db 17 LGPTPGSGSELRSAFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPAY 76
QY 61 FFSTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTWRL 120
Db 77 FFSTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTWRL 136
QY 121 HGAPHYALGAPGATFSGYLVYAD 143
Db 137 HGAPHYALGAPGATFSGYLVYAD 159

RESULT 3
ABG70384
ID ABG70384 standard; Protein; 284 AA.
XX
AC ABG70384;
XX
DT 05-NOV-2002 (first entry)
XX

XX Adipocyte complement-related C1q Tumour Necrosis Factor-like protein #2.
DE Human; NOX; NOX-associated disorder; cardiomyopathy; atherosclerosis;
XX cell signal processing; metabolic pathway modulation; metabolic disorder;
KW obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
KW Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
KW haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
KW memory defect; infertility; congenital heart defect; hair growth;
KW pigmentation disorder; endocrine disorder; respiratory disease; health;
KW gastro-intestinal disease; reproductive; neurological disease;
KW bone marrow transplantation; endocrine disease; allergy; inflammation;
KW nephrological disorder; urinary system disorder; age-related disorder;
KW neuropsychiatric disorder; EGF-related protein; SCUB1; TEN-M4;
KW adipocyte complement-related C1q tumour necrosis factor; out at first;
KW beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;
KW type Ia membrane sushi domain containing.
XX type Ia membrane-sushi domain containing.
XX Homo sapiens.
OS
XX
XX WO200257453-A2.
PN
XX
XX 25-JUL-2002.
PD
XX
XX 19-DEC-2001; 2001WO-US50331.
PF
XX
XX 19-DEC-2000; 2000US-265704P.
PR 20-DEC-2000; 2000US-257314P.
PR 02-MAY-2001; 2001US-288153P.
PR 29-MAY-2001; 2001US-294075P.
PR 24-JUL-2001; 2001US-307506P.
PR 10-AUG-2001; 2001US-311590P.
PR 10-AUG-2001; 2001US-311613P.
PR 29-AUG-2001; 2001US-315617P.
PR 14-SEP-2001; 2001US-322358P.
XX
XX (CURA-) CURAGEN CORP.
PA
XX
PI Ganggoli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
PI Stone DJ, Anderson D, Shinkets RA, Burgess CE, Zernhusen BD, Liu X;
PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
DR WPI; 2002-590744/63.
DR N-PSDB; ABS52096.
XX
XX Novel isolated NOX polypeptide useful for treating cardiomyopathy,
PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious
PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
PT cancer -
XX
XX Claim 1; Page 30; 318pp; English.
XX
XX The present invention relates to new NOX polypeptides. The invention is
CC useful for treating or preventing a NOX-associated disorder such as
CC cardiomyopathy or atherosclerosis, where the disorder is related to cell
CC signal processing and metabolic pathway modulation in a subject,
CC preferably human. The invention is also useful for treating metabolic
CC disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
CC disorders, haematopoietic disorders and various cancers. The molecules of
CC the invention are also useful for treating or preventing cirrhosis,
CC pancreatitis, learning and memory defects, infertility, congenital heart
CC defects, acne, hair growth, pigmentation disorders, endocrine disorders,
CC respiratory disease, gastro-intestinal diseases, reproductive, health,
CC neurological diseases, bone marrow transplantation, endocrine diseases,
CC allergy and inflammation, nephrological disorders, urinary system
CC disorders, neuropsychiatric disorders and age-related disorders.
CC The present amino acid sequence represents a NOX protein of the
CC invention.
XX
XX Sequence 284 AA;

Query Match 98.9%; Score 737; DB 23; Length 284;
Best Local Similarity 99.3%; Pred. No. 1.2e-75;
Matches 142; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LGFTPGSGSELRSASFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPAY 60
DB 17 LGFTPGSGSELRSASFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPAY 76
QY 61 FFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRPGARRAASQSAMQLDYGDTVWLRL 120
DB 77 FFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRPGARRAASQSAMQLDYGDTVWLRL 136
QY 121 HGAPHYALGAPGATFSGYLVYAD 143
DB 137 HGAPQYALGAPGATFSGYLVYAD 159

RESULT 4
AAB61423
ID AAB61423 standard; protein; 348 AA.
XX
XX AAB61423;
AC
XX
XX 04-APR-2001 (first entry)
DT
XX
XX Human MANGO 245 protein.
DE
XX
XX TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
XX Homo sapiens.
OS
XX
XX WO200100672-A1.
PN
XX
XX 04-JAN-2001.
PD
XX
XX 29-JUN-2000; 2000WO-US18184.
PF
XX
XX 29-JUN-1999; 99US-0342687.
PR
XX
XX (MILL-) MILLENNIUM PHARM INC.
PA
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
PI
XX
XX WPI; 2001-050127/06.
DR
XX
XX Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
XX
XX Claim 1; Fig 23; 262pp; English.
XX
XX The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
XX Sequence 348 AA;
SQ
QY Query Match 98.9%; Score 737; DB 22; Length 348;
Best Local Similarity 99.3%; Pred. No. 1.5e-75;
Matches 142; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
DB 1 LGFTPGSGSELRSASFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPAY 60
17 LGFTPGSGSELRSASFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPAY 76

QY 61 FFSFTAGKAPHKSLSVMLVRNDEVOALAFDEQRREGARRAASQSAMLQLDYGDVTWLR 120
DB 77 FFSFTAGKAPHKSLSVMLVRNDEVOALAFDEQRREGARRAASQSAMLQLDYGDVTWLR 136
QY 121 HGAPHYALGAPGATFSGYLVYAD 143
DB 137 HGAPQYALGAPGATFSGYLVYAD 159

RESULT 5

AAB61466
ID AAB61466 standard; protein; 334 AA.
XX
AC AAB61466;
XX
DT 04-APR-2001 (first entry)
XX
DE Human MANGO 245 mature protein.

XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.

XX
OS Homo sapiens.
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.

XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.

XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;

XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -

XX
PS Disclosure; Page 242; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.

XX
SQ Sequence 334 AA;
QY Query Match 98.4%; Score 733; DB 22; Length 334;
Best Local Similarity 98.6%; Pred. No. 4.1e-75;
Matches 141; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LGPTPGGSSSELRSAFSAARTTPLEGTSSEMAVTFDKVYVNIIGDPDVA TGQFRCRVPGAY 60
DB 3 LGPTPGGSSSELRGAFAARTTPLEGTSSEMAVTFDKVYVNIIGDPDVA TGQFRCRVPGAY 62

QY 61 FFSFTAGKAPHKSLSVMLVRNDEVOALAFDEQRREGARRAASQSAMLQLDYGDVTWLR 120
DB 63 FFSFTAGKAPHKSLSVMLVRNDEVOALAFDEQRREGARRAASQSAMLQLDYGDVTWLR 122

QY 121 HGAPHYALGAPGATFSGYLVYAD 143
DB 123 HGAPQYALGAPGATFSGYLVYAD 145

RESULT 6
ABG79643
ID ABG79643 standard; Protein; 329 AA.
XX
AC ABG79643;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human novel secreted protein SECP19, Incyte ID No. 931619CD1.

XX
KW Human; SECP; secreted protein; cell proliferative disorder;
KW actinic keratosis; arteriosclerosis; bursitis; hepatitis; cancer;
KW autoimmune disorder; inflammatory disorder; AIDS; asthma; allergy;
KW acquired immunodeficiency syndrome; anaemia; atopic dermatitis;
KW cardiovascular disorder; congestive heart failure; vascular tumour;
KW ischaemic heart disease; myocardial infarction; epilepsy; stroke;
KW hypertensive heart disease; neurological disorder; cerebral neoplasm;
KW Alzheimer's disease; developmental disorder; renal tubular acidosis;
KW Cushing's syndrome; Duchenne muscular dystrophy; hypothyroidism;
KW Becker muscular dystrophy.

XX
OS Homo sapiens.
XX
PN WO200262841-A2.
XX
PD 15-AUG-2002.

XX
PF 28-JAN-2002; 2002WO-US02616.

XX
PR 02-FEB-2001; 2001US-266195P.
PR 08-FEB-2001; 2001US-267924P.
PR 09-FEB-2001; 2001US-267816P.
PR 09-FEB-2001; 2001US-268112P.
PR 26-FEB-2001; 2001US-271639P.
PR 07-SEP-2001; 2001US-317818P.
PR 21-DEC-2001; 2001US-343553P.

XX
PA (INCY-) INCYTE GENOMICS INC.

XX
PI Tang TY, Yue H, Gandhi AR, Yao MG, Warren BA, Ding L, Duggan BM;
PI Xu Y, Yang J, Thangavelu K, Lal PG, Honchell CD, Walia NK, Lee S;
PI Lee EA, Richardson TW, Baughn MR, Elliott VS;

XX
DR WPI; 2002-657522/70.
DR N-PSDB; ABS64954.

XX
PT New human secreted proteins and nucleic acids useful in diagnosing,
PT treating and preventing cell proliferative, autoimmune/inflammatory,
PT cardiovascular, neurological, and developmental disorders -

XX
PS Claim 1; Page 140; 158pp; English.

XX
CC The invention relates to twenty four human secreted proteins
CC (SECP1-24), proteins 90% identical to them and active fragments of them.
CC Also included are nucleic acids encoding the SECP proteins, a recombinant
CC polynucleotide comprising a promoter sequence operably linked to the
CC nucleic acid, a cell transformed with the recombinant polynucleotide,
CC a transgenic organism comprising the recombinant polynucleotide, an
CC anti-SECP antibody, and screening for ant/agonists and modulators of
CC SECP function or expression. The SECP proteins and nucleic acids are
CC useful in the diagnosis, treatment and prevention of cell proliferative
CC (e.g. actinic keratosis, arteriosclerosis, bursitis, hepatitis or
CC cancer), autoimmune/inflammatory (e.g. AIDS (acquired immunodeficiency
CC syndrome), asthma, anaemia, allergies or atopic dermatitis),
CC cardiovascular (e.g. congestive heart failure, ischaemic heart disease,
CC myocardial infarction, hypertensive heart disease, or vascular tumours),
CC neurological (e.g. epilepsy, stroke, cerebral neoplasms, or Alzheimer's
CC disease), and developmental (e.g. renal tubular acidosis, Cushing's
CC syndrome, Duchenne and Becker muscular dystrophy, or hypothyroidism)
CC disorders. Many other diseases and disorders are listed in the
CC specification. These may also be used in assessing the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid

CC sequences of the secreted proteins. The present sequence represents a
CC SECP protein of the invention.
XX
SQ Sequence 329 AA;
Query Match 97.4%; Score 726; DB 23; Length 329;
Best Local Similarity 98.6%; Pred. No. 2.5e-74;
Matches 141; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 LGPTPGSGSELRSAFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAY 60
Db 17 LGPTPGSGSELRSAFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAY 76
OY 61 FFSFTAGKAPHKSLSVMLVRNRDEVOALAFDEQRRPGARRAASQSAMQLDYGDTWMLRL 120
Db 77 FFSFTAGKAPHKSLSVMLVRNRDEVOALAFDEQRRPGARRAASQSAMQLDYGDTWMLRL 136
OY 121 HGAPHYALGAPGATFSGYLVIYAD 143
Db 137 LGAPQYALGAPGATFSGYLVIYAD 159
RESULT 7
AAB61424
ID AAB61424 standard; protein: 329 AA.
XX
AC AAB61424;
XX
DT 04-APR-2001 (first entry)
XX
DE Monkey MANGO 245 protein.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KM autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Catarrhini sp..
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Claim 1; Fig 25; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 329 AA;
Query Match 95.0%; Score 708; DB 22; Length 329;
Best Local Similarity 95.1%; Pred. No. 2.9e-72;
Matches 136; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 LGPTPGSGSELRSAFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAY 60
Db 17 LGPTPGSGSELRSAFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAY 76
OY 61 FFSFTAGKAPHKSLSVMLVRNRDEVOALAFDEQRRPGARRAASQSAMQLDYGDTWMLRL 120
Db 77 FFSFTAGKAPHKSLSVMLVRNRDEVOALAFDEQRRPGARRAASQSAMQLDYGDTWMLRL 136
OY 121 HGAPHYALGAPGATFSGYLVIYAD 143
Db 137 LGAPQYALGAPGATFSGYLVIYAD 159
RESULT 8
AAB61488
ID AAB61488 standard; protein: 199 AA.
XX
AC AAB61488;
XX
DT 04-APR-2001 (first entry)
XX
DE Murine MANGO 245 protein.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KM autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Mus musculus.
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Claim 1; Fig 29; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 199 AA;
Query Match 95.0%; Score 707.5; DB 22; Length 199;
Best Local Similarity 95.8%; Pred. No. 1.7e-72;
Matches 137; Conservative 3; Mismatches 2; Indels 1; Gaps 1;
OY 1 LGPTPGSGSELRSAFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAY 60
Db 17 LGPA-GPGSELRSAFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDATTGQFRCRVPGAY 75
OY 61 FFSFTAGKAPHKSLSVMLVRNRDEVOALAFDEQRRPGARRAASQSAMQLDYGDTWMLRL 120
Db 76 FFSFTAGKAPHKSLSVMLVRNRDEVOALAFDQRRPGARRAASQSAMQLDYGDTWMLRL 135

QY 121 HGAPHYALGAPGATFSGYLVYAD 143
DB 136 HGAPHYALGAPGATFSGYLVYAD 158

RESULT 9
AAB61469 standard; protein; 126 AA.

AC AAB61469;

DT 04-APR-2001 (first entry)

DE Human MANGO 245 Clq domain #1.

KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KM autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.

XX Homo sapiens.

PN WO200100672-A1.

PD 04-JAN-2001.

PF 29-JUN-2000; 2000WO-US18184.

PR 29-JUN-1999; 99US-0342687.

PA (MILL-) MILLENNIUM PHARM INC.

PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;

DR WPI; 2001-050127/06.

PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -

PS Disclosure; Fig 26; 262pp; English.

CC The present invention relates to cDNAs encoding TANGO 244,

CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.

CC The nucleic acids, proteins and protein modulators are useful for

CC treating colonic disorders, inflammatory diseases, tumors,

CC renal disorders, liver disorders, lung disorders, autoimmune diseases,

CC allergic diseases, cardiovascular diseases, brain disorders,

CC degenerative diseases placental, pancreatic, skeletal and muscle

CC disorders.

XX Sequence 126 AA;

Query Match 86.7%; Score 646; DB 22; Length 126;

Best Local Similarity 99.2%; Pred. No. 1e-65; Mismatches 1; Indels 0; Gaps 0;

Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 15 AFSARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAYFFSFTAGKAPHKSL 74
DB 1 AFSARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAYFFSFTAGKAPHKSL 60

QY 75 SVMLVNRNDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTVWLRHLHGAPHYALGAPGAT 134
DB 61 SVMLVNRNDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTVWLRHLHGAPHYALGAPGAT 120

QY 135 FSGYLV 140
DB 121 FSGYLV 126

RESULT 10
AAB61489

ID AAB61489 standard; protein; 126 AA.

AC AAB61489;

DT 04-APR-2001 (first entry)

DE Murine MANGO 245 Clq domain.

KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KM autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.

XX Mus musculus.

PN WO200100672-A1.

PD 04-JAN-2001.

PF 29-JUN-2000; 2000WO-US18184.

PR 29-JUN-1999; 99US-0342687.

PA (MILL-) MILLENNIUM PHARM INC.

PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;

DR WPI; 2001-050127/06.

PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -

PS Disclosure; Page 260-261; 262pp; English.

CC The present invention relates to cDNAs encoding TANGO 244,

CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.

CC The nucleic acids, proteins and protein modulators are useful for

CC treating colonic disorders, inflammatory diseases, tumors,

CC renal disorders, liver disorders, lung disorders, autoimmune diseases,

CC allergic diseases, cardiovascular diseases, brain disorders,

CC degenerative diseases placental, pancreatic, skeletal and muscle

CC disorders.

XX Sequence 126 AA;

Query Match 85.8%; Score 639; DB 22; Length 126;

Best Local Similarity 96.8%; Pred. No. 6.5e-65; Mismatches 1; Indels 0; Gaps 0;

Matches 122; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 15 AFSARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAYFFSFTAGKAPHKSL 74
DB 1 AFSARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAYFFSFTAGKAPHKSL 60

QY 75 SVMLVNRNDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTVWLRHLHGAPHYALGAPGAT 134
DB 61 SVMLVNRNDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTVWLRHLHGAPHYALGAPGAT 120

QY 135 FSGYLV 140
DB 121 FSGYLV 126

RESULT 11
AAB61477 standard; protein; 125 AA.

AC AAB61477;

DT 04-APR-2001 (first entry)

DE Human MANGO 245 extracellular domain.

XX

KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Homo sapiens.
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Disclosure; Page 246; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 125 AA;
XX

Query Match 85.2%; Score 635; DB 22; Length 125;
Best Local Similarity 100.0%; Pred. No. 1.8e-64;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGPTPGSGSELRSAFSAARTTPLEGISEMAVTFDKVYVNIGGDFDVATGQFRCRVPGAY 60
Db 3 LGPTPGSGSELRSAFSAARTTPLEGISEMAVTFDKVYVNIGGDFDVATGQFRCRVPGAY 62

QY 61 FFSTTAGKAPHKSLSVMLVRNRDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTWTLRL 120
Db 63 FFSTTAGKAPHKSLSVMLVRNRDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTWTLRL 122

QY 121 HGA 123
Db 123 HGA 125

RESULT 12
AAB61472
ID AAB61472 standard; protein; 126 AA.
XX
AC AAB61472;
XX
DT 04-APR-2001 (first entry)
XX
DE Monkey MANGO 245 Clq domain #1.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Catarhini sp..
XX
PN WO200100672-A1.
XX

PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Disclosure; Fig 28; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 126 AA;
XX

Query Match 83.5%; Score 622; DB 22; Length 126;
Best Local Similarity 95.2%; Pred. No. 5.6e-63;
Matches 120; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 15 AFSARTTPLEGISEMAVTFDKVYVNIGGDFDVATGQFRCRVPGAYFFSTTAGKAPHKSL 74
Db 1 AFSARTTPLEGASEMAVTFDKVYVNIGGDFDAATGQFRCRVPGAYFFSTTAGKAPHKSL 60

QY 75 SVMLVRNRDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTWTLRLHGAPHYALGAPGAT 134
Db 61 SVMLVRNRDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTWTLRLHGAPHYALGAPGAT 120

QY 135 FSGYLV 140
Db 121 FSGYLV 126

RESULT 13
AAB42189
ID AAB42189 standard; Protein; 123 AA.
XX
AC AAB42189;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF1953 polypeptide sequence SEQ ID NO:3906.
XX
KW Human; open reading frame; ORFX; detection; cyrostatic; hepatotropic;
KW vulnerary; antiparinsonian; antiparkinsonian; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; proliferative disorder; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
OS Homo sapiens.
XX

XX XX WO200058473-A2.
PN
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US08621.
XX
PR 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shinkets RA, Leach M;
XX
DR WPI; 2000-602362/57.
DR N-PSDB; AAC76398.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
PS Claim 11; Page 3051-3052; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antiparasitic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatologic; immunosuppressive;
CC antihypertensive; antibacterial; antiviral; antifungal; antineoplastic;
CC antihypertensive; antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 123 AA;
SQ
Query Match 83.0%; Score 618; DB 21; Length 123;
Best Local Similarity 98.4%; Pred. No. 1.6e-62;
Matches 121; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 8 GSSELSAFAAARTTPLEGTSEMAVTFDKVVYVNIIGDFDVATGQFRCRVPAGYFFSFTAG 67
Db 1 GSSELSAFAAARTTPLEGTSEMAVTFDKVVYVNIIGDFDVATGQFRCRVPAGYFFSFTAG 60
QY 68 KAPHKSLVWLVRNDEVQALAFDEQRRPGARRAASQASAMLDYGDYDTVWLRHLGAPHYA 127
Db 61 KAPHKSPVWLVRNDEVQALAFDEQRRPGARRAASQASAMLDYGDYDTVWLRHLGAPHYA 120
QY 128 LGA 130
Db 121 LGA 123
RESULT 14
ID ABG70383 standard; Protein; 221 AA.
AC ABG70383;
XX
DT 05-NOV-2002 (first entry)

XX DE Adipocyte complement-related C1q Tumour Necrosis Factor-like protein #1.
XX
KW Human; NOVA; NOVA-associated disorder; cardiomyopathy; atherosclerosis;
KW cell signal processing; metabolic pathway modulation; metabolic disorder;
KW obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
KW Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
KW haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
KW memory defect; infertility; congenital heart defect; hair growth;
KW pigmentation disorder; endocrine disorder; respiratory disease; health;
KW gastro-intestinal disease; reproductive; neurological disease;
KW bone marrow transplantation; endocrine disease; allergy; inflammation;
KW nephrological disorder; urinary system disorder; age-related disorder;
KW neuropsychiatric disorder; EGF-related protein; SCUB1; TEN-M4;
KW adipocyte complement-related C1q tumour necrosis factor; out at first;
KW beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;
KW type Ia membrane sushi-containing domain; butyrophilin;
KW type Ia membrane-sushi domain containing.
XX
OS Homo sapiens.
XX
PN WO200257453-A2.
XX
PD 25-JUL-2002.
XX
PF 19-DEC-2001; 2001WO-US50331.
XX
PR 19-DEC-2000; 2000US-265704P.
PR 20-DEC-2000; 2000US-257314P.
PR 02-MAY-2001; 2001US-288153P.
PR 29-MAY-2001; 2001US-294075P.
PR 24-JUL-2001; 2001US-307506P.
PR 10-AUG-2001; 2001US-311590P.
PR 10-AUG-2001; 2001US-311613P.
PR 29-AUG-2001; 2001US-315617P.
PR 14-SEP-2001; 2001US-322358P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Gangoli EA, Paturajan M, Vernet CAM, Malyankar UM, Kekuda R;
PI Stone DJ, Anderson D, Shinkets RA, Burgess CE, Zerhusen BD, Liu X;
PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
XX WPI; 2002-590744/63.
DR N-PSDB; ABS52095.
XX
PT Novel isolated NOVA polypeptide useful for treating cardiomyopathy,
PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious
PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
PT cancer -
XX
PS Claim 1; Page 28; 318pp; English.
XX
XX The present invention relates to new NOVA polypeptides. The invention is
XX useful for treating or preventing a NOVA-associated disorder such as
XX cardiomyopathy or atherosclerosis, where the disorder is related to cell
XX signal processing and metabolic pathway modulation in a subject,
XX preferably human. The invention is also useful for treating metabolic
XX disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
XX disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
XX disorders, haematopoietic disorders and various cancers. The molecules of
XX the invention are also useful for treating or preventing cirrhosis,
XX pancreatitis, learning and memory defects, infertility, congenital heart
XX defects, acne, hair growth, pigmentation disorders, endocrine disorders,
XX respiratory disease, gastro-intestinal diseases, reproductive, health,
XX neurological diseases, bone marrow transplantation, endocrine diseases,
XX allergy and inflammation, nephrological disorders, urinary system
XX disorders, neuropsychiatric disorders and age-related disorders.
XX The present amino acid sequence represents a NOVA protein of the
XX invention.
SQ Sequence 221 AA;

Query Match 60.1%; Score 448; DB 23; Length 221;
Best Local Similarity 66.0%; Pred. No. 8.5e-43;
Matches 95; Conservative 10; Mismatches 31; Indels 8; Gaps 3;
OY 5 PGPSSSELSRSAFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAYFFSF 64
Db 66 PGDSSAMRSASFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDAAGVFRCLPGAYFFSF 125
OY 65 TAGKAPHKSLSVMLVRNRDEVQALAFDE--QRRPGARRAASQASAMLQLDYGDVTWMLRLH- 121
Db 126 TLGKLPRKTLNVKLMKNRDEVQAMITYDDGASRR--REMOSQSVMLALRRGDVAVMLLSHD 182
OY 122 --GAPHYALGAPGATFSGYLVYAD 143
Db 183 HDGYGAYSNHGKXITFSGFLVYPD 206

RESULT 15

AAB61468 ID AAB61468 standard; protein; 130 AA.
XX AC AAB61468;
XX DT 04-APR-2001 (first entry)
XX DE Mature monkey MANGO 245 protein.
XX KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX OS Catarrhini sp..
XX PN WO200100672-A1.
XX PD 04-JAN-2001.
XX PF 29-JUN-2000; 2000WO-US18184.
XX PR 29-JUN-1999; 99US-0342687.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX DR WPI; 2001-050127/06.
XX PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX PS Disclosure; Page 243; 262pp; English.
XX CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX SQ Sequence 130 AA;

Query Match 57.2%; Score 426.5; DB 22; Length 130;
Best Local Similarity 62.1%; Pred. No. 1.2e-40;
Matches 90; Conservative 2; Mismatches 20; Indels 33; Gaps 2;

OY 2 GPTPGPSSSELSRSAFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAYF 61
Db 1 GPAPGPSSSELSRSAFSAARTTPLEGASEMAVTFDKVYVNIIGDFDAATGQFRCRVPGAYF 60

OY 62 FSFTAGKAPHKSLSVMLVRNRDEVQALAFDEQRRPGARRAASQASAMLQLDYGDVTWMLRLH 121
Db 61 FSFTVGKAPHKSLSVMLALRR-----GDAVWLLSH 90
OY 122 --GAPHYALGAPGATFSGYLVYAD 143
Db 91 DHDGYGAYSNHGKXITFSGFLVYPD 115

Search completed: January 12, 2004, 08:15:35
Job time : 21.5499 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:15:44 ; Search time 17.4112 Seconds
(without alignments)
1655.023 Million cell updates/sec

Title: US-10-085-167-2_COPY_17_159
Perfect score: 745
Sequence: 1 LGPTPGGSSSELRSAFSAAR.....PHYALGAPGATFSGYLTVAD 143

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747907 seqs, 201509753 residues

Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------------|--------------------|
| 1 | 745 | 100.0 | 329 | US-10-085-167-2 | Sequence 2, Appli |
| 2 | 739 | 99.2 | 329 | US-10-236-055A-14 | Sequence 14, Appli |
| 3 | 706.5 | 94.8 | 326 | US-10-236-055A-16 | Sequence 16, Appli |
| 4 | 222 | 29.8 | 202 | US-10-091-458-43 | Sequence 43, Appli |
| 5 | 222 | 29.8 | 420 | US-10-236-055A-2 | Sequence 2, Appli |
| 6 | 222 | 29.8 | 746 | US-09-738-973-185 | Sequence 185, App |
| 7 | 222 | 29.8 | 746 | US-09-854-133-185 | Sequence 185, App |
| 8 | 222 | 29.8 | 746 | US-10-144-649A-185 | Sequence 185, App |
| 9 | 222 | 29.8 | 908 | US-10-309-422-10 | Sequence 10, Appli |
| 10 | 222 | 29.8 | 909 | US-10-309-422-22 | Sequence 22, Appli |
| 11 | 222 | 29.8 | 957 | US-10-309-422-14 | Sequence 14, Appli |
| 12 | 222 | 29.8 | 958 | US-10-309-422-26 | Sequence 26, Appli |
| 13 | 222 | 29.8 | 992 | US-10-309-422-18 | Sequence 18, Appli |
| 14 | 222 | 29.8 | 993 | US-10-309-422-30 | Sequence 30, Appli |
| 15 | 222 | 29.8 | 1042 | US-10-309-422-8 | Sequence 8, Appli |

| | | | | | | |
|----|-----|------|------|----|--------------------|--------------------|
| 16 | 222 | 29.8 | 1043 | 12 | US-10-309-422-20 | Sequence 20, Appli |
| 17 | 222 | 29.8 | 1091 | 12 | US-10-309-422-12 | Sequence 12, Appli |
| 18 | 222 | 29.8 | 1092 | 12 | US-10-309-422-24 | Sequence 24, Appli |
| 19 | 222 | 29.8 | 1126 | 12 | US-10-309-422-16 | Sequence 16, Appli |
| 20 | 222 | 29.8 | 1127 | 12 | US-10-309-422-28 | Sequence 28, Appli |
| 21 | 216 | 29.0 | 409 | 11 | US-10-236-055A-4 | Sequence 4, Appli |
| 22 | 200 | 26.8 | 294 | 11 | US-09-866-050A-294 | Sequence 294, App |
| 23 | 200 | 26.8 | 294 | 12 | US-10-236-055A-8 | Sequence 8, Appli |
| 24 | 197 | 26.4 | 285 | 11 | US-09-866-050A-382 | Sequence 382, App |
| 25 | 196 | 26.3 | 285 | 16 | US-10-234-000-5 | Sequence 5, Appli |
| 26 | 196 | 26.3 | 330 | 12 | US-10-236-055A-6 | Sequence 6, Appli |
| 27 | 193 | 25.9 | 146 | 12 | US-10-325-717-24 | Sequence 24, Appli |
| 28 | 193 | 25.9 | 163 | 12 | US-10-325-717-58 | Sequence 58, Appli |
| 29 | 192 | 25.8 | 146 | 12 | US-10-325-717-23 | Sequence 23, Appli |
| 30 | 192 | 25.8 | 163 | 12 | US-10-325-717-54 | Sequence 54, Appli |
| 31 | 192 | 25.8 | 163 | 12 | US-10-325-717-55 | Sequence 55, Appli |
| 32 | 192 | 25.8 | 163 | 12 | US-10-325-717-59 | Sequence 59, Appli |
| 33 | 191 | 25.6 | 144 | 12 | US-10-325-717-5 | Sequence 5, Appli |
| 34 | 191 | 25.6 | 144 | 12 | US-10-325-717-25 | Sequence 25, Appli |
| 35 | 191 | 25.6 | 145 | 12 | US-10-325-717-11 | Sequence 11, Appli |
| 36 | 191 | 25.6 | 145 | 12 | US-10-325-717-28 | Sequence 28, Appli |
| 37 | 191 | 25.6 | 157 | 12 | US-10-325-717-4 | Sequence 4, Appli |
| 38 | 191 | 25.6 | 163 | 12 | US-10-325-717-10 | Sequence 10, Appli |
| 39 | 191 | 25.6 | 163 | 12 | US-10-325-717-19 | Sequence 19, Appli |
| 40 | 191 | 25.6 | 187 | 12 | US-10-325-717-13 | Sequence 13, Appli |
| 41 | 191 | 25.6 | 187 | 12 | US-10-325-717-34 | Sequence 34, Appli |
| 42 | 191 | 25.6 | 193 | 12 | US-10-325-717-12 | Sequence 12, Appli |
| 43 | 191 | 25.6 | 193 | 12 | US-10-325-717-31 | Sequence 31, Appli |
| 44 | 191 | 25.6 | 203 | 12 | US-10-325-717-3 | Sequence 3, Appli |
| 45 | 191 | 25.6 | 203 | 12 | US-10-325-717-22 | Sequence 22, Appli |

ALIGNMENTS

RESULT 1
US-10-085-167-2
; Sequence 2, Application US/10085167
; Publication No. US20030170781A1
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: SECRETED PROTEIN ZACRP4
; FILE REFERENCE: 99-29
; CURRENT APPLICATION NUMBER: US/10/085,167
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: 60/141,928
; PRIOR FILING DATE: 1999-07-01
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 2
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-085-167-2

| | | | | |
|-----------------------|-----------------|--|-----------------|-------------|
| Query Match | 100.0%; | Score 745; | DB 12; | Length 329; |
| Best Local Similarity | 100.0%; | Pred. No. 3.6e-78; | | |
| Matches 143; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| QY | 1 | LGPTPGGSSSELRSAFSAARTTPLEGTSEMAVTFDKVYVNI | GDFDVATGQFRCRVP | GAY 60 |
| DB | 17 | LGPTPGGSSSELRSAFSAARTTPLEGTSEMAVTFDKVYVNI | GDFDVATGQFRCRVP | GAY 76 |
| QY | 61 | FFSFTAGKAPHKLSVMLVRNDEVQALAFDEQRRPGARRASOSAM | LQLDYGDYVWLR | L 120 |
| DB | 77 | FFSFTAGKAPHKLSVMLVRNDEVQALAFDEQRRPGARRASOSAM | LQLDYGDYVWLR | L 136 |
| QY | 121 | HGAPHYALGAPGATFSGYLTVAD | 143 | |
| DB | 137 | HGAPHYALGAPGATFSGYLTVAD | 159 | |

;; PRIOR APPLICATION NUMBER: 60/249,299
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/236,327
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/241,785
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/244,617
;; PRIOR FILING DATE: 2000-11-01
;; PRIOR APPLICATION NUMBER: 60/225,268
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/236,368
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/251,856
;; PRIOR FILING DATE: 2000-12-08
;; PRIOR APPLICATION NUMBER: 60/251,868
;; PRIOR FILING DATE: 2000-12-08
;; PRIOR APPLICATION NUMBER: 60/229,344
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/234,997
;; PRIOR FILING DATE: 2000-09-25
;; PRIOR APPLICATION NUMBER: 60/229,343
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/229,345
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/229,287
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/229,513
;; PRIOR FILING DATE: 2000-09-05
;; PRIOR APPLICATION NUMBER: 60/231,413
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/229,509
;; PRIOR FILING DATE: 2000-09-05
;; PRIOR APPLICATION NUMBER: 60/236,367
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/237,039
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/237,038
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/236,370
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/236,802
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/237,037
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/237,040
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/240,960
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/239,935
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: 60/239,937
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: 60/241,787
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/246,474
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/246,532
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/249,216
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,210
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/226,681
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 60/225,759
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/225,213
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/227,182
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 60/225,214
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/235,836

;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/230,438
;; PRIOR FILING DATE: 2000-09-06
;; PRIOR APPLICATION NUMBER: 60/215,135
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: 60/225,266
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/249,218
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,208
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,213
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,212
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,207
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,245
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,244
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,217
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,211
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,215
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,264
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,214
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,297
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/232,400
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/231,242
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/232,081
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/232,080
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/231,414
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/231,244
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/233,064
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/233,063
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,397
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,399
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,401
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/241,808
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,826
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,786
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,221
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/246,475
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/231,243
;; PRIOR FILING DATE: 2000-09-08

Query Match 29.8%; Score 222; DB 15; Length 202;
Best Local Similarity 41.7%; Pred. No. 1.4e-17;
Matches 58; Conservative 18; Mismatches 59; Indels 4; Gaps 4;

Db 66 PLPQOMRVAFSAARTSNLAPGTLDPVFDLLNNLGETFDLQGRFNCVPNGTYVFIFH 125
QY 66 AGK-APHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQASAMLDYDGTWMLRHGAP 124
Db 126 MLKLAENVPLVYVNLKMN-EEVLVSAYANDGAPD-HETASNHAILQLFGDQIWLRLHRGA 183
QY 125 HYALGAPGATFSGYLVYAD 143
Db 184 IYGSSWKYSTFSGYLLYQD 202

RESULT 5
US-10-236-055A-2

; Sequence 2, Application US/10236055A
; Publication No. US20030134328A1
; GENERAL INFORMATION:
; APPLICANT: Babham, Beth E.
; APPLICANT: Forsythe, Ian
; APPLICANT: Gorman, Daniel M.
; APPLICANT: Mattson, Jeanine
; APPLICANT: Moshrefi, Mehرداد
; APPLICANT: Parham, Christi
; TITLE OF INVENTION: MAMMALIAN GENES; RELATED REAGENTS
; FILE REFERENCE: DX01343K
; CURRENT APPLICATION NUMBER: US/10/236, 055A
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/317, 988
; PRIOR FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-236-055A-2

Query Match 29.8%; Score 222; DB 12; Length 420;
Best Local Similarity 41.7%; Pred. No. 3.7e-17;
Matches 58; Conservative 18; Mismatches 59; Indels 4; Gaps 4;

QY 7 PGSSSLRSAFSAARTTPL-EGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAYFFSFT 65
Db 284 PLPQOMRVAFSAARTSNLAPGTLDPVFDLLNNLGETFDLQGRFNCVPNGTYVFIFH 343
QY 66 AGK-APHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQASAMLDYDGTWMLRHGAP 124
Db 344 MLKLAENVPLVYVNLKMN-EEVLVSAYANDGAPD-HETASNHAILQLFGDQIWLRLHRGA 401
QY 125 HYALGAPGATFSGYLVYAD 143
Db 402 IYGSSWKYSTFSGYLLYQD 420

RESULT 6
US-09-738-973-185

; Sequence 185, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Flings, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indritas, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliott, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738, 973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 185
; LENGTH: 746
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-738-973-185

Query Match 29.8%; Score 222; DB 10; Length 746;
Best Local Similarity 41.7%; Pred. No. 8e-17;
Matches 58; Conservative 18; Mismatches 59; Indels 4; Gaps 4;

QY 7 PGSSSLRSAFSAARTTPL-EGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAYFFSFT 65
Db 610 PLPQOMRVAFSAARTSNLAPGTLDPVFDLLNNLGETFDLQGRFNCVPNGTYVFIFH 669
QY 66 AGK-APHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQASAMLDYDGTWMLRHGAP 124
Db 670 MLKLAENVPLVYVNLKMN-EEVLVSAYANDGAPD-HETASNHAILQLFGDQIWLRLHRGA 727
QY 125 HYALGAPGATFSGYLVYAD 143
Db 728 IYGSSWKYSTFSGYLLYQD 746

RESULT 7

US-09-854-133-185
; Sequence 185, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854, 133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 185
; LENGTH: 746
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-854-133-185

Query Match 29.8%; Score 222; DB 10; Length 746;
Best Local Similarity 41.7%; Pred. No. 8e-17;
Matches 58; Conservative 18; Mismatches 59; Indels 4; Gaps 4;

QY 7 PGSSSLRSAFSAARTTPL-EGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAYFFSFT 65
Db 610 PLPQOMRVAFSAARTSNLAPGTLDPVFDLLNNLGETFDLQGRFNCVPNGTYVFIFH 669
QY 66 AGK-APHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQASAMLDYDGTWMLRHGAP 124
Db 670 MLKLAENVPLVYVNLKMN-EEVLVSAYANDGAPD-HETASNHAILQLFGDQIWLRLHRGA 727
QY 125 HYALGAPGATFSGYLVYAD 143
Db 728 IYGSSWKYSTFSGYLLYQD 746

RESULT 8

US-10-144-649A-185
; Sequence 185, Application US/10144649A
; Publication No. US20030118599A1
; GENERAL INFORMATION:

```

; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Algate, Paul A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITL OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C11
; CURRENT APPLICATION NUMBER: US/10/144,649A
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 749
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 185
;   LENGTH: 746
;   TYPE: PRT
; ORGANISM: Homo sapien
US-10-144-649A-185

```

| | | | | |
|-----------------------|--------|------------------|--------|----------------|
| Query Match | 29.8%; | Score 222; | DB 15; | Length 746; |
| Best Local Similarity | 41.7%; | Pred. No. 8e-17; | | |
| Matches | 58; | Conservative | 18; | Mismatches 59; |
| | | | Indels | 4; |
| | | | Gaps | 4; |

QY 7 PGSELRSAFSARTPL-EGTSEMAVFDKVVNIGGDFDVATGQFRCRYEAYFFSF 65
::| | | | : : : : : : : : : :
Db 610 PLPQMRFVAFSARTSNLAPGLDQPIVFDLLNNUGETFDLQGRFNCPPVNGTVYFIH 669

QY 66 AGK-APHKSLSVMLVRNRDEVQALAEDEQRFPGARPAASQSAMLQLDYGDTIWLRLHGAP 124
 | : | | : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 670 MLKLAVNVPLLYNLMKN -EEVLVSAYNDGAPD-HETASNHAITLQLFGQDGQIMLRHGA 727

```

QY      125  HYALGAPGATFSGYLVAD  143
          |  :| ||||| :|
Db      728  IYGSSWKYSTFSGYLLYQD  746

```

```

RESULT 9
US-10-309-422-10
; Sequence 10, Application US/10309422
; Publication No. US20030139587A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20030139587A1el Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: LEX-0142-USA
; CURRENT APPLICATION NUMBER: US/10/309,422
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: US/09/798,771
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,557
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 908
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-309-422-10

```

| | | | | |
|-----------------------|--------|------------------|--------|----------------|
| Query Match | 29.8% | Score 222; | DB 12; | Length 908; |
| Best Local Similarity | 41.7%; | Pred. No. 1e-16; | | |
| Matches | 58; | Conservative | 18; | Mismatches 59; |
| | | | Indels | 4; |
| | | | Gaps | 4; |

```

Oy      7  PGSELRSAFSAARTTPL-EGTSEMAVTEDKVYVNIIGDFDVATGQFRCRYGAYFFSFT 65
      |  ::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      772 PLPQOMRVAFSARTSNLAPGTLDPVTFDLLNNLGETFDLQGRFNCVPVNGTYVFI 831

```

QY 66 AGK-APHKSLSVMLVRNRDEVQALAEDEQRFPGARRAASQSAMLQDYPGTWLR LHGAP 124
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 832 MLKLAVNVP L VYNLMKN -EEVLVSAYANDGAPD-HETASNHAITQLFQGDIQWLRLHGA 889

QY 125 HVALGARGATFSGLYVAD 143

Db 890 IYSSWKYSTFSGYLLYQD 908

RESULT 10
US-10-309-422-22

; Sequence 22, Application US/10309422
; Publication No. US20030139587A1
; GENERAL INFORMATION:

APPLICANT: walke, D. Wade
;
APPLICANT: wilganowski, Nathaniel L.
;

; APPLICANT: Turner, C. Alexander Jr.
 ; TITLE OF INVENTION: No. US20030139587
 ; FILE REFERENCE: LEX-0142-USA

; CURRENT APPLICATION NUMBER: US/10/309,422
 ; CURRENT FILING DATE: 2002-12-03
 ;

PRIOR APPLICATION NUMBER: US/09/798,771
PRIOR FILING DATE: 2001-03-02

PRIOR APPLICATION NUMBER: US 60/186,557
PRIOR FILING DATE: 2000-03-02

```

; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
;

```

```

; SEQ ID NO 22
; LENGTH: 90

```

```

/
;
; TYPE: PRT
; ORGANISM: homo sapiens
;

```

US-10-309-422-22

| | | | | |
|-----------------------|--------|------------------|--------|----------------|
| Query Match | 29.8% | Score 222; | DB 12; | Length 909; |
| Best Local Similarity | 41.7%; | Pred. No. 1e-16; | | |
| Matches | 58; | Conservative | 18; | Mismatches 59; |
| | | | Indels | 4; |
| | | | Gaps | 4; |

OY 7 PGSELRSAFSARTTPL-EGTSEMAVTPDKKYVINIGDFDVATGQFCRCRVPAGAYFFSFT 65
::| | | | | : | | : | | | | |
Db 773 PLPQOMRVAFSARTSNLAPGTLDQPIVFDDLNNUGETFDLQLGRFNCPPVNGYVFIH 832

QY 66 AGK-APHKSLSVMLVRNDEVOALAFDEQRRCARRAASQSAMLQLDYGDTWRLHGAP 124
 | : | | : : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 833 MLKLAVNVPLVYNLMKN-EEVLVSAYANDGARD-HETASNHAIIQLFQGDOIWLRLHGA 890

```
QY      125  HVALGARGATFSGYLVYAD  143
      |      :||| | | | :|
Db      891  IYGSSWKYSTFSGYLLYQD  909
```

RESULT 11
US-10-309

; Sequence 14, Application US/10309422
 ; Publication No. US20030139587A1
 ; GENERAL INFORMATION:

APPLICANT: walke, D. Wade
APPLICANT: wilganowski, Nathaniel L.

APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: NO. US20030139587A1

FILE REFERENCE: LEX-0142-USA
CURRENT APPLICATION NUMBER: US/10/309,422

; CURRENT FILING DATE: 2002-12-03
 ; PRIOR APPLICATION NUMBER: US/09/798,771

PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: US 60/186,557

PRIOR FILING DATE: 2000-03-02
NUMBER OF SEQ ID NOS: 43

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14

```

```

; LENGTH: 95
; TYPE: PRT
;

```

ORGANISM: homo sapiens
US-10-309-422-14

| | | | | |
|-----------------------|--------|--------------------|--------|----------------|
| Query Match | 29.8%; | Score 222; | DB 12; | Length 957; |
| Best Local Similarity | 41.7%; | Pred. No. 1.1e-16; | | |
| Matches | 58; | Conservative | 18; | Mismatches 59; |
| | | | Indels | 4; |
| | | | Gaps | 4; |

QY 7 PGSSELRSAFSARTTPL-EGTSEMAVTFDKVYVNIGGDFDVATGQFRCRVPGAYFFSFT 65

```
Db      821 PLPQMRVAFSAARTSNLAPGTLDPVFDLLNNLGETFDLQGRFNCVPVNGTYVFIFH 880
QY      66 AGK-APHKSLSVMLVRNRDEVQALAFDEQRRPGARRAASQSAMQLDYGDTWMLRLHGAP 124
Db      881 MLKLAVNVPLVYNLMKN-EEVLVSAYANDGAPD-HETASNHAITLQLFGDQIWLRLHRGA 938
QY      125 HYALGAPGATFSGYLVYAD 143
Db      939 IYGSSWKYSTFSGYLLYQD 957
```

RESULT 12

```
US-10-309-422-26
; Sequence 26, Application US/10309422
; Publication No. US20030139587A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20030139587A1el Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: LEX-0142-USA
; CURRENT APPLICATION NUMBER: US/10/309,422
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: US/09/798,771
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,557
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 26
; LENGTH: 958
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-309-422-26
```

```
Query Match      29.8%; Score 222; DB 12; Length 958;
Best Local Similarity 41.7%; Pred. No. 1.1e-16;
Matches 58; Conservative 18; Mismatches 59; Indels 4; Gaps 4;
```

```
QY      7 PGSSLSRASFSAARTPL-EGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAYFFSFT 65
Db      822 PLPQMRVAFSAARTSNLAPGTLDPVFDLLNNLGETFDLQGRFNCVPVNGTYVFIFH 881
QY      66 AGK-APHKSLSVMLVRNRDEVQALAFDEQRRPGARRAASQSAMQLDYGDTWMLRLHGAP 124
Db      882 MLKLAVNVPLVYNLMKN-EEVLVSAYANDGAPD-HETASNHAITLQLFGDQIWLRLHRGA 938
QY      125 HYALGAPGATFSGYLVYAD 143
Db      940 IYGSSWKYSTFSGYLLYQD 958
```

RESULT 13

```
US-10-309-422-18
; Sequence 18, Application US/10309422
; Publication No. US20030139587A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20030139587A1el Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: LEX-0142-USA
; CURRENT APPLICATION NUMBER: US/10/309,422
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: US/09/798,771
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,557
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 18
; LENGTH: 992
```

```
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-309-422-18
```

```
Query Match      29.8%; Score 222; DB 12; Length 992;
Best Local Similarity 41.7%; Pred. No. 1.2e-16;
Matches 58; Conservative 18; Mismatches 59; Indels 4; Gaps 4;
```

```
QY      7 PGSSLSRASFSAARTPL-EGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAYFFSFT 65
Db      856 PLPQMRVAFSAARTSNLAPGTLDPVFDLLNNLGETFDLQGRFNCVPVNGTYVFIFH 915
QY      66 AGK-APHKSLSVMLVRNRDEVQALAFDEQRRPGARRAASQSAMQLDYGDTWMLRLHGAP 124
Db      916 MLKLAVNVPLVYNLMKN-EEVLVSAYANDGAPD-HETASNHAITLQLFGDQIWLRLHRGA 973
QY      125 HYALGAPGATFSGYLVYAD 143
Db      974 IYGSSWKYSTFSGYLLYQD 992
```

RESULT 14

```
US-10-309-422-30
; Sequence 30, Application US/10309422
; Publication No. US20030139587A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20030139587A1el Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: LEX-0142-USA
; CURRENT APPLICATION NUMBER: US/10/309,422
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: US/09/798,771
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,557
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 30
; LENGTH: 993
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-309-422-30
```

```
Query Match      29.8%; Score 222; DB 12; Length 993;
Best Local Similarity 41.7%; Pred. No. 1.2e-16;
Matches 58; Conservative 18; Mismatches 59; Indels 4; Gaps 4;
```

```
QY      7 PGSSLSRASFSAARTPL-EGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAYFFSFT 65
Db      857 PLPQMRVAFSAARTSNLAPGTLDPVFDLLNNLGETFDLQGRFNCVPVNGTYVFIFH 916
QY      66 AGK-APHKSLSVMLVRNRDEVQALAFDEQRRPGARRAASQSAMQLDYGDTWMLRLHGAP 124
Db      917 MLKLAVNVPLVYNLMKN-EEVLVSAYANDGAPD-HETASNHAITLQLFGDQIWLRLHRGA 974
QY      125 HYALGAPGATFSGYLVYAD 143
Db      975 IYGSSWKYSTFSGYLLYQD 993
```

RESULT 15

```
US-10-309-422-8
; Sequence 8, Application US/10309422
; Publication No. US20030139587A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20030139587A1el Human Proteins and Polynucleotides Encodin
; FILE REFERENCE: LEX-0142-USA
; CURRENT APPLICATION NUMBER: US/10/309,422
```

; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: US/09/798,771
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,557
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for windows version 4.0
; SEQ ID NO 8
; LENGTH: 1042
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-309-422-8

Query Match 29.8%; Score 222; DB 12; Length 1042;
Best Local Similarity 41.7%; Pred. No. 1.2e-16;
Matches 58; Conservative 18; Mismatches 59; Indels 4; Gaps 4;

QY 7 PGSELSAFAFSARTTPL-EGTSEMAVTFDKVYVNIIGDFDVAATGQFRCRVPGAYFFSFT 65
Db 906 PLPQOMRVAFAFSARTSNLAPGTLDPQIVFDLLNNLGETFDLQGRFNCVPVNGTYVFIFH 965
QY 66 AGK-APHKSLSVMLVRNRDEVQALAFDEQRRPGARRAASQSAMLQLDYGDVTWLRHGA 124
Db 966 MLKLAVNVPLLYNLMKN-EEVLVSAYANDGAPD-HETASNHAILQLFQGDQIWLRLHGA 1023
QY 125 HYALGAPGATFSGYLVYAD 143
Db 1024 IYGSSWKYSTFSGYLLYQD 1042

Search completed: January 12, 2004, 08:33:37
Job time : 18.412 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:14 ; Search time 7.27844 Seconds
(without alignments)
831.284 Million cell updates/sec

Title: US-10-085-167-2_COPY_17_159
Perfect score: 745
Sequence: 1 LGPTPGGSSELSAFAAAR.....PHYALGAPGATFSGYLIVAD 143

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*

5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*

6: /cgn2_6/ptodata/2/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Query length | DB ID | Description |
|------------|-------|-------------|--------------|-------|--------------------------------------|
| 1 | 222 | 29.8 | 746 | 4 | US-09-370-838-185 Sequence 185, App |
| 2 | 200 | 26.8 | 294 | 3 | US-09-188-930-294 Sequence 294, App |
| 3 | 200 | 26.8 | 294 | 4 | US-09-312-283C-294 Sequence 294, App |
| 4 | 197 | 26.4 | 285 | 4 | US-09-312-283C-382 Sequence 382, App |
| 5 | 191 | 25.6 | 231 | 4 | US-09-530-423-2 Sequence 2, Appli |
| 6 | 191 | 25.6 | 244 | 3 | US-08-463-911-7 Sequence 7, Appli |
| 7 | 191 | 25.6 | 244 | 4 | US-09-140-804-3 Sequence 3, Appli |
| 8 | 191 | 25.6 | 244 | 4 | US-09-336-536-20 Sequence 20, Appli |
| 9 | 191 | 25.6 | 244 | 4 | US-09-530-423-1 Sequence 1, Appli |
| 10 | 191 | 25.6 | 244 | 4 | US-09-686-838B-3 Sequence 3, Appli |
| 11 | 191 | 25.6 | 244 | 4 | US-09-911-176B-48 Sequence 48, Appli |
| 12 | 191 | 25.6 | 244 | 4 | US-09-552-225A-3 Sequence 3, Appli |
| 13 | 191 | 25.6 | 244 | 4 | US-09-619-740-51 Sequence 51, Appli |
| 14 | 191 | 25.6 | 244 | 4 | US-09-776-976-6 Sequence 6, Appli |
| 15 | 191 | 25.6 | 244 | 4 | US-09-909-547-6 Sequence 6, Appli |
| 16 | 191 | 25.6 | 244 | 4 | US-09-569-852B-6 Sequence 6, Appli |
| 17 | 189 | 25.4 | 228 | 4 | US-09-336-536-4 Sequence 4, Appli |
| 18 | 189 | 25.4 | 243 | 3 | US-09-140-804-2 Sequence 2, Appli |
| 19 | 189 | 25.4 | 243 | 4 | US-09-336-536-3 Sequence 3, Appli |
| 20 | 189 | 25.4 | 243 | 4 | US-09-686-838B-2 Sequence 2, Appli |
| 21 | 188 | 25.2 | 259 | 4 | US-09-996-243-47 Sequence 47, Appli |
| 22 | 188 | 25.2 | 260 | 4 | US-09-489-847-198 Sequence 198, App |
| 23 | 188 | 25.2 | 287 | 4 | US-09-489-847-349 Sequence 349, App |
| 24 | 185 | 24.8 | 247 | 2 | US-08-463-911-2 Sequence 2, Appli |
| 25 | 185 | 24.8 | 247 | 3 | US-09-140-804-8 Sequence 8, Appli |
| 26 | 185 | 24.8 | 247 | 3 | US-09-118-408-3 Sequence 3, Appli |
| 27 | 185 | 24.8 | 247 | 4 | US-09-506-855-3 Sequence 3, Appli |

| | | | | | | |
|----|-------|------|-----|---|--------------------|--------------------|
| 28 | 185 | 24.8 | 247 | 4 | US-09-686-838B-8 | Sequence 8, Appli |
| 29 | 185 | 24.8 | 247 | 4 | US-09-911-176B-3 | Sequence 3, Appli |
| 30 | 185 | 24.8 | 247 | 4 | US-09-619-740-3 | Sequence 3, Appli |
| 31 | 185 | 24.8 | 247 | 4 | US-09-776-976-4 | Sequence 4, Appli |
| 32 | 185 | 24.8 | 247 | 4 | US-09-506-852-3 | Sequence 3, Appli |
| 33 | 185 | 24.8 | 247 | 4 | US-09-909-547-4 | Sequence 4, Appli |
| 34 | 179.5 | 24.1 | 243 | 3 | US-09-188-930-295 | Sequence 295, App |
| 35 | 179.5 | 24.1 | 243 | 4 | US-09-312-283C-295 | Sequence 295, App |
| 36 | 178.5 | 24.0 | 228 | 4 | US-09-336-536-11 | Sequence 11, Appli |
| 37 | 178.5 | 24.0 | 243 | 4 | US-09-336-536-10 | Sequence 10, Appli |
| 38 | 177 | 23.8 | 247 | 4 | US-09-776-976-2 | Sequence 2, Appli |
| 39 | 177 | 23.8 | 247 | 4 | US-09-909-547-2 | Sequence 2, Appli |
| 40 | 170.5 | 22.9 | 281 | 3 | US-09-118-408-2 | Sequence 2, Appli |
| 41 | 170.5 | 22.9 | 281 | 4 | US-09-506-855-2 | Sequence 2, Appli |
| 42 | 170.5 | 22.9 | 281 | 4 | US-09-911-176B-2 | Sequence 2, Appli |
| 43 | 170.5 | 22.9 | 281 | 4 | US-09-619-740-2 | Sequence 2, Appli |
| 44 | 170.5 | 22.9 | 281 | 4 | US-09-506-852-2 | Sequence 2, Appli |
| 45 | 167.5 | 22.5 | 130 | 4 | US-09-485-316A-11 | Sequence 11, Appli |

ALIGNMENTS

RESULT 1

US-09-370-838-185

Sequence 185, Application US/09370838

Patent No. 6444425

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Lodes, Michael J.

APPLICANT: Mohamath, Roadoh

APPLICANT: Secrist, Heather

TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF

FILE REFERENCE: 210121.475C1

CURRENT APPLICATION NUMBER: US/09/370,838

CURRENT FILING DATE: 1999-08-09

EARLIER APPLICATION NUMBER: US 09/285,323

EARLIER FILING DATE: 1999-04-02

NUMBER OF SEQ ID NOS: 289

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 185

LENGTH: 746

TYPE: PRT

ORGANISM: Homo sapien

US-09-370-838-185

Query Match 29.8%; Score 222; DB 4; Length 746;

Best Local Similarity 41.7%; Pred. No. 2e-18;

Matches 58; Conservative 18; Mismatches 59; Indels 4; Gaps 4;

QY 7 PGSSSELSAFAAATTPL-EGTSEMAVTPDKYVYVIGDQFDVATGQRCRVGAYFFSFT 65

DB 610 PLPQOMRVAFSAARTSNLAPGTLDPVFDLLNNLGETFDLQLGRENCPVNGTYVFIFH 669

QY 66 AGK-APHKSLVWLVNRNDEVQALAFDEQRRPGARRAASQASAMLQLDYGDVWLRHGA 124

DB 670 MLKLAIVNVPLVYVNLKMN-EEVLVSAYANDGAPD-HETASNHAILQLFOGDQIWLRLHGA 727

QY 125 HYALGAPGATFSGYLIVAD 143

DB 728 IYGSWSKYSTFSGYLLYQD 746

RESULT 2

US-09-188-930-294

Sequence 294, Application US/09188930A

Patent No. 6150502

GENERAL INFORMATION:

APPLICANT: Watson, James D.

APPLICANT: Strachan, Lorna

APPLICANT: Sleeman, Matthew

APPLICANT: Onrust, Rene

```
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 294
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Rat
; US-09-188-930-294
```

```
Query Match      26.8%; Score 200; DB 3; Length 294;
Best Local Similarity 34.9%; Pred. No. 2.7e-16;
Matches 53; Conservative 30; Mismatches 45; Indels 24; Gaps 8;
```

```
QY 5 PGP---GSSELRSAFSAA--RTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGA 59
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 147 PGPCCGSSRAKSAFSAVATKSYPRE--RLPIKFDKILMNEGGHYNASSGKFVCSVPGI 203

QY 60 YFFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDYGDVTWLR 119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 204 YFTYDITLA-NKHLAIGLVHN-GQYRIRTFD--ANTGNHDVASSGSTITLALKEGDEVWLQ 259

QY 120 LHGA-----PHYALGAPGATFSGYLVYAD 143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 260 IFYSEQNGLFYDPYWT---DSLFTGFLIYAD 287
```

RESULT 3

```
US-09-312-283C-294
; Sequence 294, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 294
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Mouse
; US-09-312-283C-294
```

```
Query Match      26.8%; Score 200; DB 4; Length 294;
Best Local Similarity 34.9%; Pred. No. 2.7e-16;
Matches 53; Conservative 30; Mismatches 45; Indels 24; Gaps 8;
```

```
QY 5 PGP---GSSELRSAFSAA--RTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGA 59
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 147 PGPCCGSSRAKSAFSAVATKSYPRE--RLPIKFDKILMNEGGHYNASSGKFVCSVPGI 203

QY 60 YFFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDYGDVTWLR 119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 204 YFTYDITLA-NKHLAIGLVHN-GQYRIRTFD--ANTGNHDVASSGSTITLALKEGDEVWLQ 259

QY 120 LHGA-----PHYALGAPGATFSGYLVYAD 143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 260 IFYSEQNGLFYDPYWT---DSLFTGFLIYAD 287
```

RESULT 4

```
US-09-312-283C-382
; Sequence 382, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 382
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Mouse
; US-09-312-283C-382
```

```
Query Match      26.4%; Score 197; DB 4; Length 285;
Best Local Similarity 34.2%; Pred. No. 6e-16;
Matches 52; Conservative 31; Mismatches 45; Indels 24; Gaps 8;
```

```
QY 5 PGP---GSSELRSAFSAA--RTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGA 59
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 138 PGPCCGSSRAKSAFSAVATKSYPRE--RLPIKFDKILMNEGGHYNASSGKFVCSVPGI 194

QY 60 YFFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDYGDVTWLR 119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 195 YFTYDITLA-NKHLAIGLVHN-GQYRIRTFD--ANTGNHDVASSGSTITLALKEGDEVWLQ 250

QY 120 LHGA-----PHYALGAPGATFSGYLVYAD 143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 251 IFYSEQNGLFYDPYWT---DSLFTGFLIYAD 278
```

RESULT 5

```
US-09-530-423-2
; Sequence 2, Application US/09530423
; Patent No. 6461821
; GENERAL INFORMATION:
; APPLICANT: Otsuka Pharmaceutical Co., Ltd.
; TITLE OF INVENTION: Smooth muscle growth inhibitory composition, a
; TITLE OF INVENTION: diagnostic method for arteriosclerosis and a kit
; TITLE OF INVENTION: therefor
; FILE REFERENCE: P98-51
; CURRENT APPLICATION NUMBER: US/09/530,423
; CURRENT FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: JP H9-297569
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Abdominal fat tissue from myoma uteri
; US-09-530-423-2
```

```
Query Match      25.6%; Score 191; DB 4; Length 231;
Best Local Similarity 31.8%; Pred. No. 2.4e-15;
Matches 47; Conservative 29; Mismatches 54; Indels 18; Gaps 6;
```

```
QY 5 PGPSSSELRSAFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAYFFSF 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 91 PGEAGAYVYRSASFVGLETYVT-IPNMPIRFTKIFYNQONHXYDGSSTGKFHCNIPGLYYFAY 149

QY 65 TAGKAPH-----KSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDYGDVTWLR 119
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 150 -----HITVYMKDVKVSIFK-KDKAMLFTYDQYQENNVDO-ASGSVLHLHLEVGDQVWLQ 201
```

OY 120 LHGAPH---YALGAPGATFSGYLVYAD 143
 DB 202 VYGEGERNGLYADNDNDSTFTGFLLYHD 229

RESULT 6

US-08-463-911-7
 ; Sequence 7, Application US/08463911
 ; Patent No. 5869330
 ; GENERAL INFORMATION:
 ; APPLICANT: Scherer, Philipp E.
 ; APPLICANT: Lodish, Harvey F.
 ; TITLE OF INVENTION: A NOVEL SERUM PROTEIN PRODUCED
 ; TITLE OF INVENTION: EXCLUSIVELY IN ADIPOCYTES
 ; NUMBER OF SEQUENCES: 7
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
 ; STREET: Two Militia Drive
 ; CITY: Lexington
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02173
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/463,911
 ; FILING DATE:
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Granahan, Patricia
 ; REGISTRATION NUMBER: 32,227
 ; REFERENCE/DOCKET NUMBER: WHI95-05
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 861-6240
 ; TELEFAX: (617) 861-9540
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 244 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-463-911-7

Query Match 25.6%; Score 191; DB 2; Length 244;
 Best Local Similarity 31.8%; Pred. No. 2.6e-15;
 Matches 47; Conservative 29; Mismatches 54; Indels 18; Gaps 6;

OY 5 PGPGSSSLRSASFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAYFFSF 64
 DB 104 PGEAGVYVRSASFVGLTETVT-IPNMPIRFTKIFYNQNNHYDGTGKFHCNIPGLYFFAY 162
 OY 65 TAGKAPH-----KSLSVMLVRNDEVQALAFDEQRRPGARRAASQASAMLQLDYGDYTWLR 119
 DB 163 -----HITVYMKDVKVSLEK-KDKAMLFTYDQYQENNVQ-ASGSVLLHLEVGDDQVWLQ 214
 OY 120 LHGAPH---YALGAPGATFSGYLVYAD 143
 DB 215 VYGEGERNGLYADNDNDSTFTGFLLYHD 242

RESULT 7

US-09-140-804-3
 ; Sequence 3, Application US/09140804
 ; Patent No. 6197930
 ; GENERAL INFORMATION:
 ; APPLICANT: Shepard, Paul O.
 ; APPLICANT: Humes, Jacqueline M.
 ; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
 ; FILE REFERENCE: 97-49

; CURRENT APPLICATION NUMBER: US/09/140,804
 ; CURRENT FILING DATE: 1998-08-26
 ; EARLIER APPLICATION NUMBER: 60/056,983
 ; EARLIER FILING DATE: 1997-08-26
 ; NUMBER OF SEQ ID NOS: 47
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 3
 ; LENGTH: 244
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-140-804-3

Query Match 25.6%; Score 191; DB 3; Length 244;
 Best Local Similarity 31.8%; Pred. No. 2.6e-15;
 Matches 47; Conservative 29; Mismatches 54; Indels 18; Gaps 6;

OY 5 PGPGSSSLRSASFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAYFFSF 64
 DB 104 PGEAGVYVRSASFVGLTETVT-IPNMPIRFTKIFYNQNNHYDGTGKFHCNIPGLYFFAY 162
 OY 65 TAGKAPH-----KSLSVMLVRNDEVQALAFDEQRRPGARRAASQASAMLQLDYGDYTWLR 119
 DB 163 -----HITVYMKDVKVSLEK-KDKAMLFTYDQYQENNVQ-ASGSVLLHLEVGDDQVWLQ 214
 OY 120 LHGAPH---YALGAPGATFSGYLVYAD 143
 DB 215 VYGEGERNGLYADNDNDSTFTGFLLYHD 242

RESULT 8

US-09-336-536-20
 ; Sequence 20, Application US/09336536
 ; Patent No. 6406884
 ; GENERAL INFORMATION:
 ; APPLICANT: Leiby, K.
 ; APPLICANT: McKay, C.
 ; APPLICANT: Bossone, S.
 ; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
 ; FILE REFERENCE: 7853-144
 ; CURRENT APPLICATION NUMBER: US/09/336,536
 ; CURRENT FILING DATE: 1999-06-18
 ; NUMBER OF SEQ ID NOS: 75
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 20
 ; LENGTH: 244
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-336-536-20

Query Match 25.6%; Score 191; DB 4; Length 244;
 Best Local Similarity 31.8%; Pred. No. 2.6e-15;
 Matches 47; Conservative 29; Mismatches 54; Indels 18; Gaps 6;

OY 5 PGPGSSSLRSASFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAYFFSF 64
 DB 104 PGEAGVYVRSASFVGLTETVT-IPNMPIRFTKIFYNQNNHYDGTGKFHCNIPGLYFFAY 162
 OY 65 TAGKAPH-----KSLSVMLVRNDEVQALAFDEQRRPGARRAASQASAMLQLDYGDYTWLR 119
 DB 163 -----HITVYMKDVKVSLEK-KDKAMLFTYDQYQENNVQ-ASGSVLLHLEVGDDQVWLQ 214
 OY 120 LHGAPH---YALGAPGATFSGYLVYAD 143
 DB 215 VYGEGERNGLYADNDNDSTFTGFLLYHD 242

RESULT 9

US-09-530-423-1
 ; Sequence 1, Application US/09530423
 ; Patent No. 6461821
 ; GENERAL INFORMATION:
 ; APPLICANT: Otsuka Pharmaceutical Co., Ltd.
 ; TITLE OF INVENTION: Smooth muscle growth inhibitory composition, a


```

; TITLE OF INVENTION: diagnostic method for arteriosclerosis and a kit
; TITLE OF INVENTION: therefor
; FILE REFERENCE: P98-51
; CURRENT APPLICATION NUMBER: US/09/530,423
; CURRENT FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: JP H9-297569
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Abdominal fat tissue from myoma uteri
US-09-530-423-1

```

[illegible]

```

RESULT 10
US-09-686-838B-3
; Sequence 3, Application US/09686838B
; Patent No. 6482612
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Humes, Jacqueline M.
; TITLE OF INVENTION: Adipocyte-Specific Protein Homologs
; FILE REFERENCE: 97-49D1
; CURRENT APPLICATION NUMBER: US/09/686,838B
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/140,804
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: US 60/056,983
; PRIOR FILING DATE: 1997-08-26
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-686-838B-3

```

[illegible]

RESULT 11

```

US-09-911-176B-48
; Sequence 48, Application US/09911176B
; Patent No. 6518403
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: ANTIBODIES THAT BIND AN
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOG
; FILE REFERENCE: 97-30D1
; CURRENT APPLICATION NUMBER: US/09/911,176B
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: 09/118,408
; PRIOR FILING DATE: 1998-07-17
; PRIOR APPLICATION NUMBER: 60/053,154
; PRIOR FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-911-176B-48

```

[illegible]

```

RESULT 12
US-09-552-225A-3
; Sequence 3, Application US/09552225A
; Patent No. 6521233
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Bishop, Paul
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOG ZACRP3
; FILE REFERENCE: 99-09
; CURRENT APPLICATION NUMBER: US/09/552,225A
; CURRENT FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/130,199
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-552-225A-3

```

[illegible]

QY

THIS PAGE BLANK (USPTO)

F:33-116/Domain: collagenous, triple helix <COL>
F:123-249/Domain: complement C1q carboxyl-terminal homology <C1Q>
F:28/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F:31/Disulfide bonds: interchain (to chain A-26) #status experimental
F:35,38,41,53,56,65,83,86,101,104,107/Modified site: 4-hydroxyproline (Pro) #status expe
F:59,62,77,92,98,110/Modified site: 5-hydroxylysine (Lys) #status experimental
F:59,62,98,110/Binding site: carbohydrate (Lys) (covalent) #status experimental

Query Match 11.1%; Score 182; DB 1; Length 253;
Best Local Similarity 29.1%; Pred. No. 1e-07;
Matches 48; Conservative 32; Mismatches 61; Indels 24; Gaps 5;

QY 149 PARGPAPPEP-----RSAPSAARTSLVSGDAGPPRHQPLAFDTEFNIG 195
Db PKGPGAPGAPGPKGESGDYKATQKIAFSATRTINV-----PLRRDQTRFDHVTNNM 150
QY 196 GDFDAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMTYDDGASRRREM 254
Db 151 NNYEPRSGKFTCKVPGLYFTYHASS--RGNLCVNLMRGRERACKVTFCDYAYNTFQVT 208
QY 255 SQSVMLALRRGDAVWLLSHDHDGYGAYSNHGKYITFSGFLVYPDL 299
Db 209 TGGWVLKLEOGENVFLQATDKN--SLLMEGANSIFSGFLLFPDM 251

RESULT 5
S23298
collagen alpha 1(VIII) chain - chicken
C:Species: Gallus gallus (chicken)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S23298
R:Ninomiya, Y.; Castagnola, P.; Gerecke, D.; Gordon, M.K.; Jacenko, O.; LuValle, P.; McC
maguchi, N.; Olsen, B.R.
in Extracellular Matrix Genes, Sandell L.J. and Boyd C.D., eds., pp.79-114, Academic Pre
A:Title: The molecular biology of collagens with short triple-helical domains.
A:Reference number: S22243
A:Accession: S23298
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-744 <NIN>
C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
F:617-743/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 10.7%; Score 175; DB 1; Length 744;
Best Local Similarity 25.3%; Pred. No. 1.4e-06;
Matches 56; Conservative 28; Mismatches 77; Indels 60; Gaps 8;

QY 120 LHGAPHYALGAPGATFSGYLVYADADADAPA-RGPPAPPEPRS----- 161
Db 540 LHGPP-----GKPGA-----LGPGQPGLPGPPGPPGPPAVMPPTPAQGEYLPDMG 589
QY 162 -----AFSAARTSLVSGDAGP-----PRHQPLAFDTEFNIGG 196
Db 590 LGIDGVKTPHAYAAKK-----GKNGPAYEMPAFTAELTAPFPVGAPIKFDRLLYNGRQ 644
QY 197 DFDAAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMTYDDGASRRREM 256
Db 645 NYNPQTGIFTCEVPGVYFYAHV-HCKGGNVWVALLFKNNEPVM-YTYDEYKKGFLDQASG 702
QY 257 SVMLALRRGDAVWLLSHDHDGYGAYSNHGKYITFSGFLVYP 297
Db 703 SAVLLLRPGDRVFLQNPSEQAAGLYAGQYVHSSFSGYLLYP 743

RESULT 6
A34246
collagen alpha 1(VIII) chain precursor - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A34246
R:Yamaguchi, N.; Benya, P.D.; van der Rest, M.; Ninomiya, Y.
J. Biol. Chem. 264, 16022-16029, 1989
A:Title: The cloning and sequencing of alpha1(VIII) collagen cDNAs demonstrate that type

omains similar to those of type X collagen.

A:Reference number: A34246; MUID:89380199; PMID:2476437

A:Accession: A34246

A:Molecule type: mRNA

A:Residues: 1-744 <YAM>

A:Cross-references: GB:J05042; NID:g164895; PIDN:AAA31204.1; PID:g164896

C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology

F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-744/Product: collagen alpha 1(VIII) chain #status predicted <MAT>

F:21-117/Region: amino-terminal nonhelical

F:118-571/Region: interrupted helical

F:572-744/Region: carboxyl-terminal nonhelical

F:617-743/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 10.6%; Score 173; DB 1; Length 744;
Best Local Similarity 25.3%; Pred. No. 2.1e-06;
Matches 56; Conservative 27; Mismatches 78; Indels 60; Gaps 8;

QY 120 LHGAPHYALGAPGATFSGYLVYADADADAPA-RGPPAPPEPRS----- 161
Db 540 LHGPP-----GKPGA-----LGPGQPGLPGPPGPPGPPAVMPPTPAQGEYLPDMG 589
QY 162 -----AFSAARTSLVSGDAGP-----PRHQPLAFDTEFNIGG 196
Db 590 LGIDGVKTPHAYAAKK-----GKNGPAYEMPAFTAELTAPFPVGAPIKFDRLLYNGRQ 644
QY 197 DFDAAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMTYDDGASRRREM 256
Db 645 NYNPQTGIFTCEVPGVYFYAHV-HCKGGNVWVALLFKNNEPVM-YTYDEYKKGFLDQASG 702
QY 257 SVMLALRRGDAVWLLSHDHDGYGAYSNHGKYITFSGFLVYP 297
Db 703 SAVLLLRPGDRVFLQNPSEQAAGLYAGQYVHSSFSGYLLYP 743

RESULT 7

C1HUQC

complement subcomponent C1q chain C precursor - human

N:Alternate names: complement subcomponent C1q gamma chain

C:Species: Homo sapiens (man)

C:Date: 22-May-1981 #sequence_revision 31-May-1996 #text_change 22-May-1998

C:Accession: S14351; A03207

R:Sellier, G.C.; Blake, D.J.; Reid, K.B.M.

Biochem. J. 274, 481-490, 1991

A:Title: Characterization and organization of the genes encoding the A-, B- and C-chains

A:Reference number: S14350; MUID:91174759; PMID:1706597

A:Accession: S14351

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-245 <SEL>

R:Reid, K.B.M.

Biochem. J. 179, 367-371, 1979

A:Title: Complete amino acid sequences of the three collagen-like regions present in sub

A:Reference number: A90304; MUID:80020137; PMID:486087

A:Accession: A03207

A:Molecule type: protein

A:Residues: 29-56, 'P', 58-65, 'K', 67-71, 'P', 73-83, 'K', 85-86, 'D', 88-89, 'N', 91-122 <REI>

C:Comment: The first component of complement is a calcium-dependent complex of the three

ivation of C1r (enzyme), C1s (proenzyme), and the other eight components of complement.

C:Comment: The C1q subcomponent is composed of nine subunits, six of which are disulfide-

dimers of the C chain. Equimolar amounts of the A, B, and C chains are found after reduc

C:Genetics:

A:Gene: GDB:C1QG

A:Cross-references: GDB:128132; OMIM:120575

A:Map position: 1p36.3-1p34.1

A:Introns: 60/3

C:Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal homc

C:Keywords: complement pathway; glycoprotein; homodimer; hydroxylysine; hydroxyproline; l

F:1-28/Domain: signal sequence #status predicted <SIG>

F:29-245/Product: complement subcomponent C1q chain B #status predicted <MAT>

F:31-114/Domain: collagenous, triple helix <COL>

F:121-244/Domain: complement C1q carboxyl-terminal homology <C1Q>

F:32/Disulfide bonds: interchain #status experimental

F:36,39,42,45,54,63,81,93,96,99,105/Modified site: 4-hydroxyproline (Pro) #status experi
F:57,72,75/Modified site: 5-hydroxylysine (Lys) #status experimental
F:75/Binding site: carboxylate (Lys) (covalent) #status experimental

Query Match 10.5%; Score 172; DB 1; Length 245;
Best Local Similarity 23.8%; Pred. No. 6.4e-07;
Matches 67; Conservative 33; Mismatches 84; Indels 98; Gaps 13;

QY 70 PHKSLVLMVRNDEVQALAFDEQRRPGARRAASQSAMLQLDYGDVTWLRLHGAPHYALG 129
DB 9 PHLGLKLL-----LALLALRGQANTGC-----YG-----IFGMFGLP-G 43
QY 130 APGATFSGY--LVYADADADAPA-----RGPAPP----- 157
DB 44 APGK--DGYDGLPXPKEPGIRPAIPGIRPGKQKGPGLPGHPKNGPMGPMPGVPGP 101
QY 158 -----EP-----RSAFSAARTSLVSGDAGFGRHQP-----LAPDTEFVNIGG 196
DB 102 MGIRGEPEGEGRYKQKQSVFTVTR-----QTHQPPAPNSLIRFNAVLTNPOG 149
QY 197 DFDAAAGVFRCLPGAYFFSFTLGKLPKRTLSTKLMKNRDEVQAMITYDDGASRRREMOSQ 256
DB 150 DYDTSTGKFTCKVPGLYFVYHAS---HTANLCVLLYRSGVKVTFPGHSTKTNQVNSG 205
QY 257 SVMALRRGDVAVMLSHDHDGYGAYSNHGKYITPSGFLVYPD 298
DB 206 GVLRLQVGEEVWLA VNDY--YDMVGIGGSDSVPSGFLLPD 245

RESULT 8

collagen alpha 1(VIII) chain precursor - human
C/Species: Homo sapiens (man)
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Aug-1999
C/Accession: S15435
R/Muragaki, Y.; Mattei, M.G.; Yamaguchi, N.; Olsen, B.R.; Ninomiya, Y.
Eur. J. Biochem. 197, 615-622, 1991
A/Title: The complete primary structure of the human alpha-1(VIII) chain and assignment
A/Reference number: S15435; MUID:91231001; PMID:2029894
A/Accession: S15435
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-744 <MUR>
A/Cross-references: EMBL:X57527; NID:930081; PIDN:CAA40748.1; PID:930082
C/Genetics:
A/Gene: GDB:COL8A1
A/Cross-references: GDB:128104; OMIM:120251
A/Map position: 3q11.1-3q13.2
C/Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-744/Product: collagen alpha 1(VIII) chain #status predicted <MAT>
F:21-117/Region: amino-terminal nonhelical
F:118-571/Region: interrupted helical
F:572-744/Region: carboxyl-terminal nonhelical
F:617-743/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 10.4%; Score 170; DB 2; Length 744;
Best Local Similarity 25.9%; Pred. No. 3.7e-06;
Matches 56; Conservative 27; Mismatches 83; Indels 50; Gaps 9;

QY 120 LHGAPHYALGAPGATSGYLIVYADADADAPA-RGPPAPPEPRS----- 161
DB 540 LHGPP---GKPGA-----LPGQGPGPLGPPGPPGPPAVMPPTPPPGGYLPDMG 589
QY 162 -AFSAARTSLVSGDAGP--GPRHQLAFDTE---FVNIGG-----DFDAA 201
DB 590 LGIDGVKPRPHATGAKKGKNGGPAVEPAFTALTAFPFPVGVKFNKLLYNGRQNYNPQ 649
QY 202 AGVFRCLPGAYFFSFTLGKLPKRTLSTKLMKNRDEVQAMITYDDGASRRREMOSQSVMLA 261
DB 650 TGI FTCEVPGVYFAHYH-HCKGGMVWVALFKNNBPMV-YTYDEVKKGFLDQASGSAYLL 707
QY 262 LRRGDVAVMLSHDHDGYGAYSNHGKYITPSGFLVYP 297

DB 708 LRPGRVFLQMPSEQAAGLYAGQYVHSSFGYLLYP 743

RESULT 9

S23297
collagen alpha 1(X) chain precursor - chicken
N/Alternate names: type X collagen
C/Species: Gallus gallus (chicken)
C/Date: 07-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 13-Aug-1999
C/Accession: S23297; A31896; S65594; S77711; I50218
R/Ninomiya, Y.; Castagnola, P.; Gerecke, D.; Gordon, M.K.; Jacenko, O.; LuValle, P.; McCa
maguchi, N.; Olsen, B.R.
in Extracellular Matrix Genes, Sandell L.J. and Boyd C.D., eds., pp.79-114, Academic Pres
A/Title: The molecular biology of collagens with short triple-helical domains.
A/Reference number: S22243
A/Accession: S23297
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-674 <NIN>
R/LuValle, P.; Ninomiya, Y.; Rosenblum, N.D.; Olsen, B.R.
J. Biol. Chem. 263, 18378-18385, 1988
A/Title: The type X collagen gene. Intron sequences split the 5'-untranslated region and
A/Reference number: A31896; MUID:89054019; PMID:2461368
A/Accession: A31896
A/Molecule type: mRNA
A/Residues: 1-75 <LUV>
R/Ninomiya, Y.; Gordon, M.; van der Rest, M.; Schmid, T.; Linsenmayer, T.; Olsen, B.R.
J. Biol. Chem. 261, 5041-5050, 1986
A/Title: The developmentally regulated type X collagen gene contains a long open reading
A/Reference number: I50218; MUID:86168227; PMID:3082876
A/Accession: S65594
A/Molecule type: DNA
A/Residues: 'T', '9', 'D', '11-12', 'EDQMKLYLFTW', '30-31', 'TCKSGRAFTYWLQNMADLVSSHT', '48-89', 'L', '5
629, 'QCAVLSLISWRTIKCGSSCOIQNPVSIPLNMFILLSGVSYLLKSNIPLTMS' <NIN1>
A/Cross-references: EMBL:M13496; NID:G211639; PIDN:AAA48736.1; PID:G211700
A/Accession: S77711
A/Molecule type: protein
A/Residues: 104-112, 'X', '114-117;453-466 <NIN2>
C/Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
C/Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer; hydroxyproline;
F:1-18/Domain: signal sequence #status predicted <SIG>
F:547-673/Domain: complement C1q carboxyl-terminal homology <C1Q>
F:453,456/Modified site: hydroxyproline (Pro) #status experimental
F:611/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 10.3%; Score 169.5; DB 2; Length 674;
Best Local Similarity 25.3%; Pred. No. 3.6e-06;
Matches 58; Conservative 21; Mismatches 75; Indels 75; Gaps 7;

QY 129 GAPGATFSGYLVYADADADAPA-----RGPAPPEPR----- 160
DB 454 GAPG-----AKGEAGAPGLPGPAGIATKGLRGPMGPPGPPGKNGSGEPLGPPG 504
QY 161 -----SAFSAARTSLVSG-----SDAGPGRHQL 185
DB 505 PPGPGQSTIPGKYVKGESRELSCMSFMKAGANQALTGMPVSAFTVILSKAYPGAT-VPI 563
QY 186 AFDTEFVNIGDFDAAAGVFRCLPGAYFFSFTLGKLPKRTLSTKLMKNRDEVQAMITYDD 245
DB 564 KFDKILYNRQHYDPTGTGIFTCRIPGLYFYSYHV-HAKGTNVWVALYNGSPVM-YTYDE 621
QY 246 GASRRREMOSQSVMLALRRGDVAVMLSHDHDGYGAYSNHGKYITPSGFL 294
DB 622 YQGYLDQASGSAVIDLMENDQVWLQLPNSBSNGLYSSEYVHSSFSGFL 670

RESULT 10

S49158
complement protein C1q beta chain precursor - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 20-Aug-1999
C/Accession: S49158

R;Schwaeble, W.; Petry, F.; Loos, M.
submitted to the EMBL Data Library, March 1993
A;Description: cDNA cloning and expression of the mRNA encoding the B-chain of rat C1q.
A;Reference number: S49158
A;Accession: S49158
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-253 <SCH>
A;Cross-references: EMBL:X71127; NID:g510191; PIDN:CAA50440.1; PID:g510192
C;Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom
F;121-249/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 10.3%; Score 168.5; DB 2; Length 253;
Best Local Similarity 29.5%; Pred. No. 1.3e-06;
Matches 46; Conservative 36; Mismatches 59; Indels 15; Gaps 7;

QY 2 GTPGP-----GSSELR-----AFSAARTTPTLEGTSEMAVTFDKVYVNIIGDFDVATGQF 52
DB 100 GP-PGPRGPKGSGDYKATQKVAFSALRTVNSALRPNOAIRFEKVTITVNDNYEPRSGKF 158
QY 53 RCRVPGAYFFSFTGKAPHKSLSVMLVRNDEVQ-ALAFDEQRRPGARRAASQSAMLQLD 111
DB 159 TCKVPGLYFTTHASSRGNLCVNIVRGRDRMQKVLTFCDYAO-NTFOYTTGGVTLKE 217
QY 112 YGDTVMLRLHGAPHYALGAPGAT--FSGYLYTADAD 145
DB 218 QEEVHLQATD-KNSLLGVEGANSITGFLFPDMD 252

RESULT 11
S31216
collagen alpha 1(X) chain precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text_change 13-Aug-1999
C;Accession: S31216; S28807; S22215; S30127; I48299; S26397; S31830
R;Kong, R.Y.C.; Kwan, K.M.; Lau, E.T.; Thomas, J.T.; Boot-Handford, R.P.; Grant, M.E.; C
Eur. J. Biochem. 213, 99-111, 1993
A;Title: Intron-exon structure, alternative use of promoter and expression of the mouse
A;Reference number: S31216; MUID:93238750; PMID:8477738
A;Accession: S31216
A;Molecule type: DNA
A;Residues: 1-680 <KON>
A;Cross-references: EMBL:Z21610; NID:g49793; PIDN:CAA79736.1; PID:g49794
R;Elima, K.; Berola, I.; Rosati, R.; Metsaeranta, M.; Garofalo, S.; Peraelae, M.; de Cro
Biochem. J. 289, 247-253, 1993
A;Title: The mouse collagen X gene: complete nucleotide sequence, exon structure and exp
A;Reference number: S28807; MUID:93143676; PMID:8424763
A;Accession: S28807
A;Molecule type: DNA
A;Residues: 1-285, 'A', 287-680 <ELI>
A;Cross-references: EMBL:X67348; NID:g50480; PIDN:CAA47763.1; PID:g50481
R;Elima, K.; Metsaeranta, M.; Kallio, J.; Peraelae, M.; Berola, I.; Garofalo, S.; de Cro
Biochim. Biophys. Acta 1130, 78-80, 1992
A;Title: Specific hybridization probes for mouse alpha-2(IX) and alpha-1(X) collagen mRN
A;Reference number: S22215; MUID:92182017; PMID:1543751
A;Accession: S22215
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 385-450, 'K', 452-627 <ELA>
A;Cross-references: EMBL:X63013; NID:g49795; PIDN:CAA44741.1; PID:g49796
R;Appe, S.S.; Olsen, B.R.
Matrix 13, 165-179, 1993
A;Title: Characterization of the mouse type X collagen gene.
A;Reference number: S30127; MUID:93261348; PMID:8492743
A;Accession: S30127
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-12, 'F', 14-26, 'S', 28-247, 'L', 249-285, 'A', 287-305, 'F', 307-416, 'S', 418-499, 'L
R;Appe, S.S.; Seldin, M.F.; Hayashi, M.; Olsen, B.R.
Eur. J. Biochem. 206, 217-224, 1992
A;Title: Cloning of the human and mouse type X collagen genes and mapping of the mouse t
A;Reference number: I48299; MUID:92267014; PMID:1587271
A;Accession: I48299

A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 52-247, 'L', 249-285, 'A', 287-305, 'F', 307-416, 'S', 418-499, 'L', 501-566, 'C', 568, 'I
A;Cross-references: EMBL:X65121; NID:g50482; PIDN:CAA46237.1; PID:g667031
R;Summers, T.A.; Irwin, M.H.; Mayne, R.; Balian, G.
J. Biol. Chem. 263, 581-587, 1988
A;Title: Monoclonal antibodies to type X collagen. Biosynthetic studies using an antibody
A;Reference number: S26397; MUID:88087150; PMID:2826450
A;Accession: S26397
A;Molecule type: protein
A;Residues: 'SDGYFSQ', 24-26, 'KQ' <SUM>
C;Genetics:
A;Gene: Col10a-1
A;Map position: 10
A;Intons: 51/3
C;Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
C;Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-680/Product: collagen alpha 1(X) chain #status predicted <MAT>
F;553-679/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 10.3%; Score 168.5; DB 2; Length 680;
Best Local Similarity 25.7%; Pred. No. 4.4e-06;
Matches 56; Conservative 23; Mismatches 82; Indels 57; Gaps 6;

QY 129 GAPGATFSGYLVYADADADAPARGPAPPEPRSAFSAARTSLVGS DAGPRHQ----- 183
DB 470 GNPGA--PGPAGIATKLGNGPT-GPPGPBPGR--GHSGEPLPGPPGPBPQAVMPD 523
QY 184 -----PLAFDTFVNIGDFD 199
DB 524 GFKAGQRPRLSGMPLVSANHGVTGMPVSAFTVILSKAYPAVGAPIPFDEILYNRQHYD 583
QY 200 AAAGVFRCLPGAYFFSFTLGKLPKRLSVKLMKNRDEVQAMTYDDGASRRREMOSQSV 259
DB 584 PRSGIFTKIPGITYFFSYHV-HVKGIHVWVGLYKNGTPTM-YTYDEYSKGYLDQASGSAI 641
QY 260 LALRRGDAVWLSHDHDGYGAYSNHGKRYITFSGFLVYP 297
DB 642 MELTENDQVWLQLPNAESNGLYSSEYVHSSFSGLVAP 679

RESULT 12
CGHVD
collagen alpha 1(X) chain precursor - human
N;Alternate names: procollagen alpha 1(X) chain
C;Species: Homo sapiens (man)
C;Date: 22-Nov-1993 #sequence revision 03-Nov-1995 #text_change 22-Jun-1999
C;Accession: S26396; S30086; S15826; S18249; A43901; I51870; S21856
R;Reichenberger, E.; Beier, F.; LuValle, P.; Olsen, B.R.; von der Mark, K.; Berling, W.
FEBS Lett. 311, 305-310, 1992
A;Title: Genomic organization and full-length cDNA sequence of human collagen X.
A;Reference number: S26396; MUID:93012005; PMID:1397333
A;Accession: S26396
A;Molecule type: DNA
A;Residues: 1-680 <REL>
A;Cross-references: EMBL:X68952; EMBL:X72578; EMBL:X72579; EMBL:X72580; GB:S47714; GB:S4
R;Appe, S.S.
submitted to the EMBL Data Library, March 1992
A;Reference number: S30085
A;Accession: S30086
A;Molecule type: DNA
A;Residues: 'TIPFYGWVCWVCL', 52-680 <APT>
A;Cross-references: EMBL:X65120; NID:g23129
A;Note: the initial difference is probably due to translation of an intronic sequence
R;Appe, S.; Mattei, M.G.; Olsen, B.R.
FEBS Lett. 282, 393-396, 1991
A;Title: Cloning of human alpha-1(X) collagen DNA and localization of the COL10A1 gene to
A;Reference number: S15826; MUID:91243838; PMID:2037056
A;Accession: S15826
A;Molecule type: DNA
A;Residues: 561-647, 'G', 649-666 <AP2>
A;Cross-references: EMBL:X58879; NID:g30013; PIDN:CAA41686.1; PID:g30014

R;Thomas, J.T.; Cresswell, C.J.; Rash, B.; Nicolai, H.; Jones, T.; Solomon, E.; Grant, M.
Biochem. J. 280, 617-623, 1991
A;Title: The human collagen X gene. Complete primary translated sequence and chromosomal
A;Reference number: S18249; PMID:1764025
A;Accession: S18249

A:Molecule type: DNA
A;Residues: 1-26,'T',28-680 <THO>
A;Cross-references: EMBL:X60382; NID:g30094; PIDN:CAA42933.1; PID:g30095
A;Note: the sequence from Fig. 3 is inconsistent with that from Fig. 2 in having 532-Ala
R;Reicheberger, E.; Aigner, T.; von der Mark, K.; Stoss, H.; Bertling, W.
Dev. Biol. 148, 562-572, 1991
A;Title: In situ hybridization studies on the expression of type X collagen in fetal hum
A;Reference number: A43901; MUID:92077285; PMID:1743401
A;Accession: A43901

A:Molecule type: mRNA
A;Residues: 547-656 <RE2>
A;Cross-references: GB:M74050; GB:D57494; NID:g339884; PIDN:AAA61221.1; PID:g553796
A;Note: sequence extracted from NCBI backbone (NCBIN:69012, NCBIP:69014)
R;Wallis, G.A.; Rash, B.; Sweetman, W.A.; Thomas, J.T.; Super, M.; Evans, G.; Grant, M.E
Am. J. Hum. Genet. 54, 169-178, 1994
A;Title: Amino acid substitutions of conserved residues in the carboxyl-terminal domain
pe Schmid.

A;Reference number: I51870; MUID:94136476; PMID:8304336
A;Accession: I51870
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A;Residues: 520-597,'D',599-680 <WAL>
A;Cross-references: GB:S68531; NID:g545180; PIDN:AAC60615.1; PID:g545181
A;Note: mutant sequence from patient with metaphyseal chondrodysplasia type Schmid
A;Note: a second mutant sequence with 614-Pro is also described
C;Comment: Prolines and lysines at the third position of the tripeptide repeating unit
ed and subsequently O-glycosylated.
C;Genetics:

A;Gene: GDB:COL10A1
A;Cross-references: GDB:I28635; OMIM:120110
A;Map position: 6q21-6q22
A;Introns: 52/1
A;Note: a defect in this gene may cause Schmid metaphyseal chondrodysplasia
C;Complex: type X collagen may be a homotrimer
C;Function:
A>Description: structural component of extracellular fibrous polymer specifically and tr
be important for skeletogenesis
C;Superfamily: collagen alpha 1(VIII) chain; complement Clg carboxyl-terminal homology
C;Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer; hydroxyllysine;
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-680/Product: collagen alpha 1(X) chain #status predicted <MAT>
F;19-56/Domain: amino-terminal nonhelical #status predicted <NC2>
F;57-519/Region: interrupted helical
F;520-680/Domain: amino-terminal nonhelical #status predicted <NC1>
F;553-679/Domain: complement Clg carboxyl-terminal homology <Clg>
F;617/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.2%; Score 166.5; DB 1; Length 680;
Best Local Similarity 25.7%; Pred. No. 6.4e-06;
Matches 56; Conservative 22; Mismatches 83; Indels 57; Gaps 6;

OY 129 GAPGATFSGYLVADADAPARGPAPPPEPRSAARTSLVGS DAGPGRHQ----- 183
 | :||| : ||||| : |||||
Db 470 GSFGP -PGPAGIATKGINOPT-GPPGPGRP--GHSGEPGLP GP PGPAQAVME 523

OY 184 -----PLAFDTFEVNIGGFDF 199
 :||| :||| :|||
Db 524 GFIKAGQRPSLSGTPLVSANGVGM PVS AFVTILSKAYPALIGTP IPFDKLIN RQH YD 583

OY 200 AAAGVFRCLPGA VFFSFTLGKLPRTLSVKLMKNRDDEVQA MIYDDGASRRREMOSQSVM 259
 |:|||||:|||::: :||| :|||
Db 584 PRTGIFTC QIPGIY FSTHV-HVGKT HWVWG LYN GT PVM-YTYDEYTKGYLDQASGSAI 641

OY 260 LA LRGD AVWL SHDH DYGAYS NHGKYIT FS GLVP 297
 :||| :||| :|||
Db 642 IDLTENDQ VW LQ LPNAESNGLYSS EYVH SS FGFLVA P 679

RESULT 13
S23779
collagen alpha 1(VIII) chain - mouse
C/Species: Mus musculus (house mouse)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C/Accession: S23779
R/Muragaki, Y.; Shiota, C.; Inoue, M.; Ooshima, A.; Olsen, B.R.; Ninomiya, Y.
Eur. J. Biochem. 207, 895-902, 1992
A/Title: Alpha-1(VIII)-collagen gene transcripts encode a short-chain collagen polypeptide
A/Reference number: S23779; NUID:92362626; PMID:1499564
A/Accession: S23779
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-743 <MUR>
A/Cross-References: EMBL:X66976; NID:g50493; PIDD:CAA47387.1; PID:g1359953
C/Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
F;616-742/Domain: complement C1q carboxyl-terminal homology <C1Q>

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 10.1%; | Score 165.5; | DB 1; | Length 743; |
| Best Local Similarity | 25.8%; | Pred. No. 8.7e-06; | | |
| Matches 56; | Conservative 23; | Mismatches 85; | Indels 53; | Gaps 6; |

QY 120 LHGAPHYALGAPGATFSGYLTVADADADAPARGPPAPPEPRSAFSAARTRSLVGS---DA 176
||| | ||| :
Db 540 LHGPP---GKPGA-----LGPGQGPGLPGRPPGPPGPPRAVMPTRSPQGEYLPDM 587

QY 177 GPG-----PRHQPLAFDTEFVNIGGDEDA 200

Db 588 GLGIDGVKTPHAYAGKKGGPAYEMPAFTAELTVFPFVGAPVKFDKLLYNGRQNYNP 647

| | | | | |
|----|-------|-----|--|-----|
| QY | : | 201 | AAGVFCRLPGAYFFSFETLGLKLPRTLKSVKLMKNRDEVQAMITYDGDASRRREMOSQSVM. | 260 |
| | : : | | : : : : | |
| Db | : | 648 | QTGIFTCEVPGVYFYFAHYH-HCKGNVWVALFKN-NPMMYTYDEYKKGFILDQASGS AVL | 705 |

QY . 261 ALRRGDAVLLSHDDGDGYAASNHGKIYTFSGFLVYP 297
||| : | : || : : ||| : ||
Db 706 LLRPGDQVFLQNPFEOAGLYAGQYVHSSFSFGYLVP 742

RESULT 14

collagen alpha 1(X) chain precursor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 21-Nov-1993 #sequence_revision 23-Feb-1996 #text_change 13-Aug-1999
C;Accession: S13301
R;Thomas, J.T.; Kwan, A.P.L.; Grant, M.E.; Boot-Handford, R.P.

Biochem. J. 273, 141-148, 1991

A;Title: Isolation of cDNAs encoding the complete sequence of bovine type X collagen. Evi
A;Reference number: S13301; MUID:9113131; PMID:1703407

A;Accession: S13301

A;Molecule type: mRNA

A;Residues: 1-674 <THO>

A;Cross-references: EMBL:X53556; NID:g263; PIDN:CAA37624.1; PID:g264

C; Genetics:

A;Gene: COL10A1

C;Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
C;Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-674/Product: collagen alpha 1(X) chain #status predicted <MAT>
F;547-673/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 10.0%; Score 163.5; DB 2; Length 674;

Best Local Similarity 26.8%; Pred. No. 1.1e-05;
Matches 57; Conservative 18; Mismatches 87; Indels 51; Gaps 6;

```
OY      128 LGAPGATFSGYLVTADADADAFAEGPPAPPEPRSAFSAARTSLVGS DAGPGRHQ---- 183
          ||| | : ||||| |||
Db      469 VGTDP--PGPAGIAVKGLNPT-GPPGP PGP R---GNAGEPLGPLSPGP PGP GPQVALP 522
```

QY 184 -----PLAFDTEFVNIGGDFDAAGV 204

Db 523 EDFVKAGQRPFSVANSQGV T G M P V S A F T V I L S K A Y P A I G T P I P F D K I L Y N K Q Q H Y D P R T G I 582

QY 205 FRCRLPGAYFFSF^TLGLPKR^TLSVKLMKNRDEVAQAM^IYDDGASRRREMOQSVMALARR 264
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 583 FTCKIPGIYYFSYHI-HVKGTHAWAGLYKNGT^PVMI-YTYDEYIKGYLDQA^SGSAVIDL^TE 640

QY 265 GDAVWL^LSHDHDGYGAYSNHGKYITFSGFLVP 297
||| |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 641 NDQVWLQLPNAGSNGLYSPPEYVHSSFGFLVAP 673

RESULT 15

cerebellin-like glycoprotein - rat
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 31-Mar-2000
 C/Accession: A60032
 R/Wada, C.; Ohtani, H:
 Brain Res. Mol. Brain Res. 9, 71-77, 1991
 A/Title: Molecular cloning of rat cerebellin-like protein cDNA which encodes a novel mem
 A/Reference number: A60032; MUID:91203483; PMID:1850079
 A/Accession: A60032
 A/Status: not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 1-224 <WAD>
 C/Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom
 C/Keywords: glycoprotein; transmembrane protein
 F/32-49/Domain: transmembrane #status predicted <TMM>
 F/50-224/Domain: extracellular #status predicted <EXT>
 F/94-223/Domain: complement C1q carboxyl-terminal homology <C1Q>
 F/53,110/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.98; Score 163; DB 2; Length 224;

Best Local Similarity 38.0%; Pred. No. 3.2e-06;

Matches 52; Conservative 15; Mismatches 62; Indels 8; Gaps 4;

QY 162 AFSARTSLVGS D A G P G P R H Q P L A F D T E F V N I G G D F D A A G V F R C R L P G A Y F F S F T L G K 221

Db 94 AFSA--TRSTNHEPSEMNRMTIYFDQVLNIGNHFDLASSIFVAPRKGISFSFHVVK 151

222 L-PRKTLVKLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRRGDAVLLSHDHGCGA 280

Db 152 VYNRQTIQVSLMQNGYPVISAFAGD-QDVTREASNGVLLMEREDKVLKLERGNLMGG 210

QY 281 YSNHGKYITFSGFLVYP 297

Db 211 W----KYSTFSGFLVFP 223

Search completed: January 12, 2004, 08:16:43
Job time : 20.1178 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:09 ; Search time 11.2455 Seconds
(without alignments)
1308.910 Million cell updates/sec

Title: US-10-085-167-2_COPY_17_329

Perfect score: 1639

Sequence: 1 LGPTPGGSSSLRSAFSAAR.....LVYPLAPAPPGIGASSELL 313

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 1639 | 100.0 | 329 | 1 CQT4_HUMAN | Q9bxj3 homo sapien |
| 2 | 215 | 13.1 | 281 | 1 CQT1_HUMAN | Q9bxj1 homo sapien |
| 3 | 213 | 13.0 | 243 | 1 CQT5_HUMAN | Q9bxj0 homo sapien |
| 4 | 201 | 12.3 | 285 | 1 CQT2_HUMAN | Q9bxj5 homo sapien |
| 5 | 200 | 12.2 | 278 | 1 CQT6_HUMAN | Q9bxj9 homo sapien |
| 6 | 199 | 12.1 | 258 | 1 C1RF_MOUSE | O88992 mus musculu |
| 7 | 192 | 11.7 | 244 | 1 APM1_HUMAN | Q15848 homo sapien |
| 8 | 192 | 11.7 | 258 | 1 C1RF_HUMAN | O75973 homo sapien |
| 9 | 188 | 11.5 | 247 | 1 APM1_MOUSE | Q60994 mus musculu |
| 10 | 186.5 | 11.4 | 255 | 1 GLIC_MOUSE | Q9esn4 mus musculu |
| 11 | 184 | 11.2 | 215 | 1 HP25_TAMSI | Q06576 tamias sibi |
| 12 | 182 | 11.1 | 251 | 1 C1QB_HUMAN | P02746 homo sapien |
| 13 | 181 | 11.0 | 289 | 1 CQT7_HUMAN | Q9bxj2 homo sapien |
| 14 | 173 | 10.6 | 744 | 1 CA18_RABIT | P14282 Oryctolagus |
| 15 | 171 | 10.4 | 419 | 1 COLE_LEPMA | P98085 lepomis mac |
| 16 | 170 | 10.4 | 245 | 1 C1QC_MOUSE | P31721 rattus norv |
| 17 | 168.5 | 10.3 | 253 | 1 C1QB_RAT | Q05306 mus musculu |
| 18 | 168.5 | 10.3 | 680 | 1 CA1A_MOUSE | P08125 gallus gall |
| 19 | 167.5 | 10.2 | 744 | 1 CA1A_CHICK | P27658 homo sapien |
| 20 | 167 | 10.2 | 680 | 1 CA1A_HUMAN | Q03692 homo sapien |
| 21 | 166.5 | 10.2 | 743 | 1 CA18_MOUSE | Q00780 mus musculu |
| 22 | 166.5 | 10.2 | 246 | 1 CA1A_HUMAN | Q9bxj4 homo sapien |
| 23 | 163.5 | 10.0 | 674 | 1 CA1A_BOVIN | P23206 bos taurus |
| 24 | 163.5 | 10.0 | 224 | 1 CERL_RAT | P98087 rattus norv |
| 25 | 163 | 9.9 | 224 | 1 C1QC_MOUSE | Q02105 mus musculu |
| 26 | 160.5 | 9.8 | 246 | 1 C1QB_MOUSE | P14106 mus musculu |
| 27 | 160.5 | 9.8 | 253 | 1 C1QA_MOUSE | P98086 mus musculu |
| 28 | 159 | 9.7 | 245 | 1 C1QA_MOUSE | P83371 oncorhynchu |
| 29 | 158.5 | 9.7 | 508 | 1 C1QA_MOUSE | P23435 homo sapien |
| 30 | 156.5 | 9.5 | 193 | 1 CERB_HUMAN | O9r171 mus musculu |
| 31 | 156.5 | 9.5 | 193 | 1 CERB_MOUSE | Q06575 tamias sibi |
| 32 | 150 | 9.2 | 196 | 1 HP20_TAMSI | P02745 homo sapien |
| 33 | 150 | 9.2 | 245 | 1 C1QA_HUMAN | |

| | | | | | |
|----|-------|-----|------|--------------|--------------------|
| 34 | 144.5 | 8.8 | 215 | 1 HP27_TAMSI | Q06577 tamias sibi |
| 35 | 144 | 8.8 | 1228 | 1 ECM_HUMAN | Q13201 homo sapien |
| 36 | 139.5 | 8.5 | 635 | 1 CA28_HUMAN | P25067 homo sapien |
| 37 | 139 | 8.5 | 201 | 1 CERL_HUMAN | Q9ntu7 homo sapien |
| 38 | 135.5 | 8.3 | 170 | 1 CA28_MOUSE | P25318 mus musculu |
| 39 | 98.5 | 6.0 | 265 | 1 Y176_HUMAN | Q14681 homo sapien |
| 40 | 96.5 | 5.9 | 476 | 1 LEU2_STRCO | O86534 streptomyc |
| 41 | 95.5 | 5.8 | 493 | 1 YKA7_CAEEL | P34258 caenorhabdi |
| 42 | 94 | 5.7 | 1045 | 1 GUNB_CELFI | P26325 cellulomona |
| 43 | 91 | 5.6 | 673 | 1 ZDH8_HUMAN | Q9ulc8 homo sapien |
| 44 | 91 | 5.6 | 2505 | 1 CCAA_HUMAN | O00555 homo sapien |
| 45 | 90.5 | 5.5 | 4289 | 1 TENX_HUMAN | P22105 homo sapien |

ALIGNMENTS

RESULT 1

ID CQT4_HUMAN STANDARD; PRT; 329 AA.
AC Q9BXJ3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Complement-clq tumor necrosis factor-related protein 4 precursor.
GN C1QTNF4 OR CTRP4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Holloway J.L., Lok S.;
RT "Homo sapiens complement-clq tumor necrosis factor-related protein.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 2 C1Q domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AF329838; AAK17962.1; -.
DR Genew; HGNC:14346; C1QTNF4.
DR InterPro; IPR001073; C1q.
DR Pfam; PF00386; C1q; 2.
DR SMART; SM00110; C1Q; 2.
DR PROSITE; PS01113; C1Q; 2.
KW Repeat; Signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 329 COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT DOMAIN 23 159 RELATED PROTEIN 4.
FT DOMAIN 170 314 C1Q 1.
SQ SEQUENCE 329 AA; 35265 MW; 331CTDBF26036915 CRC64;

Query Match 100.0%; Score 1639; DB 1; Length 329;
Best Local Similarity 100.0%; Pred. No. 1.9e-126;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| QY | 1 | LGPTPGGSSSLRSAFSAARTTPLEGTSEMAVTFDKVYVNIIGDFVATGQFCRVP | 60 |
| DB | 17 | LGPTPGGSSSLRSAFSAARTTPLEGTSEMAVTFDKVYVNIIGDFVATGQFCRVP | 76 |
| QY | 61 | FFSFTAGKAPKSLVMLVNRNDEVQALAFDEQRPGARRAASQASMLQLDYGD | 120 |
| DB | 77 | FFSFTAGKAPKSLVMLVNRNDEVQALAFDEQRPGARRAASQASMLQLDYGD | 136 |
| QY | 121 | HGAPHYALGAPGATFSGYLVYADADADAPARGPAPPEPRSAFSAARTSLV | 180 |


```

Db      137 HGAPHYALGARGATFSGVLVYADADADAPARGPPAPPEPRSAFSAARTSLVGSNAGPGR 196
QY      181 RHQPLAFDTEFVNIGSDFDAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQA 240
        |||
Db      197 RHQPLAFDTEFVNIGSDFDAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQA 256
QY      241 MIYDDGASRRREMOSQSVMLALRGGDAVWLSHDHGCGAYSNHGKXITFSGLVYPPDLA 300
        |||
Db      257 MIYDDGASRRREMOSQSVMLALRGGDAVWLSHDHGCGAYSNHGKXITFSGLVYPPDLA 316
QY      301 PAAPPLGASELL 313
        |||
Db      317 PAAPPLGASELL 329

RESULT 2
COT1_HUMAN STANDARD; PRT; 281 AA.
ID      COT1_HUMAN
AC      Q9BXJ1; Q96NF2; Q9GZR4;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Complement-c1q tumor necrosis factor-related protein 1 precursor
DE      (G protein coupled receptor interacting protein) (GIP).
GN      CIQTNF1 OR CTRP1.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Innamorati G., Le Gouill C., Whang I., Birnbaumer M.;
RT      "GIP, a putative GPCR interacting protein.";
RL      Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Sheppard P.O.;
RT      "Homo sapiens complement-c1q tumor necrosis factor-related protein.";
RL      Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Skin;
RX      MEDLINE=22388257; PubMed=12477932;
RA      Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA      Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA      Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA      Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA      Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schultz J., Myers R.M.,
RA      Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smailus D.E.,
RA      Schnerch A., Schein J.E., Jones S.J.M., Maira M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length
RT      human and mouse cDNA sequences.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN      [4]
RP      SEQUENCE OF 83-281 FROM N.A.
RA      Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA      Hotuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J.,
RA      Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA      Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA      Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA      Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA      Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA      Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;

```

```

RT      "NEDO human cDNA sequencing project.";
RL      Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC      -1- SIMILARITY: Contains 1 collagenous domain.
CC      -1- SIMILARITY: Contains 1 CIQ domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AJ272138; CAC20425.1; ALT_INIT.
DR      EMBL; AF232905; AAG44303.1; ALT_INIT.
DR      EMBL; AF329840; AAK17964.1; -.
DR      EMBL; BC021553; AAH21553.1; -.
DR      EMBL; AK055541; BAB70947.1; -.
DR      Genew; HGNC:14324; CIQTNF1.
DR      InterPro; IPR001073; CIQ.
DR      InterPro; IPR000087; Collagen.
DR      Pfam; PF00386; CIQ; 1.
DR      Pfam; PF01391; Collagen; 1.
DR      SMART; SM00110; CIQ; 1.
DR      PROSITE; PS01113; CIQ; FALSE_NEG.
KW      Collagen; Signal.
FT      SIGNAL 1 25
FT      CHAIN 26 281
FT      FT
FT      DOMAIN 99 140
FT      FT
FT      DOMAIN 147 281
FT      FT
FT      CONFLICT 22 22 L -> P (IN REF. 1).
FT      CONFLICT 241 241 R -> Q (IN REF. 1).
SQ      SEQUENCE 281 AA; 31743 MW; 49E248CB88ACFB7C CRC64;

Query Match 13.1%; Score 215; DB 1; Length 281;
Best Local Similarity 30.2%; Pred. No. 1.7e-10;
Matches 77; Conservative 26; Mismatches 86; Indels 66; Gaps 11;

QY      53 RCRVPGAYFFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASOSAMLQLDY 112
        |||
Db      75 RCCDPGTSMYPATA--VQINITILKEKGR-----GDRG-----LQKX 113
QY      113 GDT--VWLRLLGAP--HYALGARGATFSGVLVYADADADAPARGPPAPPEPRSAFSAAR 167
        |||
Db      114 GKTGSAGARGHTGPKQKSGMGAPEGCKSH--YA-----AFSVGR 152
QY      168 TRSLVGSNAGPGRH-----QPLAFDTEFVNIGSDFDAAGVFRCLPGAYFFSFTLGKL 222
        :
Db      153 KK-----PMSNHYQTIVFTFEVNLVDHFNMTGKFCYVPGLYFFSLNVHTW 202
QY      223 PRKTLVKLMKNRDEVQAMIVDDGASRRREMOSQSVMLALRGGDAVWLSHDHGCGAYS 282
        :
Db      203 NQKETYLHIMKNEEVVILFAQVG--DRSIMGOSLMLRLREQDVWVRLYKGERENALF 260
QY      283 NH--GKYITFSGLV 295
        :
Db      261 SEELDTYITFSGLV 275

RESULT 3
COT5_HUMAN STANDARD; PRT; 243 AA.
ID      COT5_HUMAN
AC      Q9BXJ0; Q9UFX4;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Complement-c1q tumor necrosis factor-related protein 5 precursor.
GN      CIQTNF5 OR CTRP5.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxID=9606;

```

[1]
RN SEQUENCE FROM N.A.
RP Sheppard P.O., Humes J.M.;
RA "Homo sapiens complement-c1q tumor necrosis factor-related protein.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE OF 25-243 FROM N.A.
RP OTTENWAEELDER B., OBERMAIER B., MEWES H.-W., GASSENHUBER J.,
RC TISSUE=Uterus;
RA Wiemann S.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C1Q domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
DR EMBL; AF329841; AAK17965.1; -.
DR EMBL; AL110261; CAB53702.1; -.
DR PIR; T14782; T14782.
DR GeneW; HGNC:14344; C1QTNF5.
DR InterPro; IPR001073; C1Q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; FALSE_NEG.
KW Collagen; Signal.
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 16 243 COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT DOMAIN 30 95 RELATED PROTEIN 5.
FT DOMAIN 97 243 COLLAGEN-LIKE.
SQ SEQUENCE 243 AA; 25298 MW; 7CCDA65CDA7EB784 CRC64;

Query Match 13.0%; Score 213; DB 1; Length 243;
Best Local Similarity 34.9%; Pred. No. 2.1e-10;
Matches 66; Conservative 18; Mismatches 83; Indels 22; Gaps 6;

QY 122 GAPHYALGAPGATFSG---YLIVYADAD---ADAPARGPPAPP-----EPRSAFSAATR 169
DB 57 GAP----GAPGEKEGEGRPLGPGRGDPGRGEAGPAGPTGPAGECSVPPRSASFASKRSE 112

QY 170 SLVGS DAGPGRHQPLAFDTEFVNIGGDFDAAGVFRCLPGAYFFSFITLGLPKRTLSV 229
DB 113 SRV----PPPSDAPLPFDRLVINEQGHYDAVTGKFTCQPGVYFA-VHATVYRASLQF 166

QY 230 KLMKNRDEVQAMIVDDGASRRREMOSQSVMLALRRGDAVWLLSHDHDGYGAYSNHGKIYT 289
DB 167 DLVKNGESIASFFQFGGWPKPASLSGAMVRLPEPDQVWVQGVGDYIGIYASIKTDST 226

QY 290 FSGFLVYPD 298
DB 227 FSGFLVYSD 235

RESULT 4
COT2_HUMAN
ID COT2_HUMAN STANDARD; PRT; 285 AA.
AC O9BXJ5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement-c1q tumor necrosis factor-related protein 2 precursor.
GN C1QTNF2 OR CTRP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP Piddington C.S., Bishop P.;
RA "Homo sapiens complement-c1q tumor necrosis factor-related protein.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP TISSUE=Muscle;
RC MEDLINE=22388257; PubMed=12477932;
RX Strausberg R.L., Pelngold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hejeh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stepleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C1Q domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
DR EMBL; AF329836; AAK17960.1; -.
DR EMBL; BC011699; AAH11699.1; -.
DR GeneW; HGNC:14325; C1QTNF2.
DR InterPro; IPR001073; C1Q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 2.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Collagen; Signal.
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 16 285 COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT DOMAIN 40 141 RELATED PROTEIN 2.
FT DOMAIN 143 285 COLLAGEN-LIKE.
SQ SEQUENCE 285 AA; 29952 MW; 7E31FF9868D4EDFA CRC64;

Query Match 12.3%; Score 201; DB 1; Length 285;
Best Local Similarity 34.4%; Pred. No. 2.4e-09;
Matches 53; Conservative 30; Mismatches 47; Indels 24; Gaps 8;

QY 5 GGP---GSSELSAFSA--RTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGA 59
DB 138 GGPSCGSGHTKSAFSAVATKSYPR--RLPIKFDKILMNEGHYNASSGKFCVGPGI 194

QY 60 YFFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQASMLQLDYGDTWLR 119
DB 195 YFTYDITLA-NKHLAIGLVHN-GQYRIRTFD--ANTGNHDVAGSSTILALKQGDDEVWLQ 250

QY 120 LHGA-----PHYALGAPGATFSGYLIVYADAD 145

Db 251 IFYSEONGLFYDPYWT-----DSLFTGFLIYADQD 280

RESULT 5

COT6_HUMAN STANDARD; PRT; 278 AA.
 ID COT6_HUMAN
 AC Q9BXI9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Complement-clq tumor necrosis factor-related protein 6 precursor.
 GN C1QTNF6 OR CTRP6.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Piddington C.S., Sheppard P.O.;
 RT "Homo sapiens complement-clq tumor necrosis factor-related protein.";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Jackson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- SIMILARITY: Contains 1 collagenous domain.
 CC -1- SIMILARITY: Contains 1 C1Q domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF329842; AAK17966.1; -;
 DR EMBL; BC020551; AAH20551.1; -;
 DR Genew; HGNC:14343; C1QTNF6.
 DR InterPro; IPR001073; C1q.
 DR InterPro; IPR000087; Collagen.
 DR Pfam; PF00386; C1q; 1.
 DR Pfam; PF01391; Collagen; 1.
 DR PRINTS; PR00007; COMPLEMENTC1Q.
 DR PROSITE; PS01113; C1Q; FALSE_NEG.
 DR Collagen; Signal.
 KW SIGNAL
 FT CHAIN 1 46 POTENTIAL.
 FT 47 278 COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
 FT DOMAIN 97 138 RELATED PROTEIN 6.
 FT 139 278 COLLAGEN-LIKE.
 FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CONFLICT 21 21 G -> V (IN REF. 2).
 SQ SEQUENCE 278 AA; 30861 MW; 27A82CA863F23D47 CRC64;

Query Match 12.2%; Score 200; DB 1; Length 278;
 Best Local Similarity 31.3%; Pred. No. 2.8e-09;
 Matches 87; Conservative 21; Mismatches 106; Indels 64; Gaps 12;

QY 32 VTFDKVVYVNIIGDQDVATGQFRC-----RYGVAFYFSTAGKAPHKSYVLMVRNDEVQ 86
 Db 46 LTFDRA-----VASGCQRCCDSEPLDPAHVSSASSSGRPHALPEI----- 86
 QY 87 ALAFDEQRRPGARRAASQSAQLQDYDPTVWLRHGAAPHY-----ALGAPGATFSGYLIV 141
 Db 87 -----RPYINITILKGDKGDCPDM---GLPGYMGREGPGQEGPQGS----- 125
 QY 142 ADADADAPARGP-APPEPR-SAFSAARTSLVGSADGPGPRHQPPLAFDTEFVNIIGDFD 199
 Db 126 --KDGKMGSGPGAPCQKRFFAFSVGKRTAL-----HSGEDFQTLFPERFVNLDGCFD 177
 QY 200 AAAGVFCRLPGAYFESFTLGKLPKRTLSVYLMKNRDEVQAMITYDDGASRRREMOSQSV 259
 Db 178 MATGQFAAPLRGIYFESLNVHSMWYKETYVHIMNQKE--AVILYAQPSERSIMQSQSV 235
 QY 260 LALRGDAVWLISHDHDGYGA-YSN-HGKYITFSGFLV 295
 Db 236 LDLAYGDRVWVRLEFKRQRENAIYSDNDFYITFSGHLI 273

RESULT 6

CIRF_MOUSE STANDARD; PRT; 258 AA.
 ID CIRF_MOUSE
 AC O88992;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE C1q-related factor precursor.
 GN C1QRF OR CRF.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99097006; PubMed=9878755;
 RA Berube N.G., Swanson X.H., Bertram M.J., Kittle J.D., Didenko V.,
 RA Baeklin D.S., Smith J.R., Pereira-Smith O.M.;
 RT "Cloning and characterization of CRF, a novel C1q-related factor,
 RT expressed in areas of the brain involved in motor function.";
 RL Brain Res. Mol. Brain Res. 63:233-240(1999).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAINSTEM. MORE ABUNDANT IN AREAS
 CC OF THE NERVOUS SYSTEM INVOLVED IN MOTOR FUNCTION, SUCH AS THE
 CC PURKINJE CELLS OF THE CEREBELLUM, THE ACCESSORY OLIVARY NUCLEUS,
 CC THE PONS AND THE RED NUCLEUS.
 CC -1- SIMILARITY: Contains 1 collagenous domain.
 CC -1- SIMILARITY: Contains 1 C1Q domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF095155; AAC64187.1; -;
 DR MGD; MGI:1344400; C1qrf.
 DR GO; GO:0005737; C:cytoplasm; IDA.
 DR InterPro; IPR001073; C1q.
 DR InterPro; IPR000087; Collagen.
 DR Pfam; PF00386; C1q; 1.
 DR Pfam; PF01391; Collagen; 1.
 DR PRINTS; PR00007; COMPLEMENTC1Q.
 DR SMART; SM00110; C1Q; 1.

[illegible]

```

RESULT 7
APM1_HUMAN
ID APM1_HUMAN STANDARD; PRT; 244 AA.
AC Q15848;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Adiponectin precursor (30 kDa adipocyte complement-related protein)
DE (ACRP30) (Adipose most abundant gene transcript 1) (apm-1) (Gelatin-
DE binding protein).
GN APM1 OR ACRP30 OR GBP28.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adipose tissue;
RX MEDLINE=96224171; PubMed=8619847;
RA Maeda K., Okubo K., Shimomura I., Funahashi T., Matsuzawa Y.,
RA Matsubara K.;
RT "cDNA cloning and expression of a novel adipose specific collagen-like
RT factor, apm1 (Adipose Most abundant Gene transcript 1).";
RL Biochem. Biophys. Res. Commun. 221:286-289(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99196984; PubMed=10095105;
RA Saito K., Tobe T., Minoshima S., Asakawa S., Sumiya J., Yoda M.,
RA Nakano Y., Shimizu N., Tomita M.;
RT "Organization of the gene for gelatin-binding protein (GBP28).";
RL Gene 229:67-73(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99333693; PubMed=10403784;
RA Schaeffler A., Orso E., Palitzsch K.D., Buechler C., Drobnik W.,
RA Fuerst A., Schoelmerich J., Schmitz G.;
RT "The human apm-1, an adipocyte-specific gene linked to the family of
RT TNF's and to genes expressed in activated T cells, is mapped to
RT chromosome 1q21.3-q23, a susceptibility locus identified for familial
RT combined hyperlipidemia (FCH).";

```

RL Biochem. Biophys. Res. Commun. 260:416-425(1999).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=20417747; PubMed=10961870;
 RA Yokota T., Oritani K., Takahashi I., Ishikawa J., Matsuyama A.,
 RA Ouchi N., Kihara S., Funahashi T., Tenner A.J., Tomiyama Y.,
 RA Matsuzawa Y.;
 RT "Adiponectin, a new member of the family of soluble defense collagens,
 RT negatively regulates the growth of myelomonocytic progenitors and the
 RT functions of macrophages.";
 RL Blood 96:1723-1732(2000).
 RN [5]
 RP CHARACTERIZATION.
 RX MEDLINE=20440368; PubMed=10982546;
 RA Ouchi N., Kihara S., Arita Y., Okamoto Y., Maeda K., Kuriyama H.,
 RA Hotta K., Nishida M., Takahashi M., Muraguchi M., Ohmoto Y.,
 RA Nakamura T., Yamashita S., Funahashi T., Matsuzawa Y.;
 RT "Adiponectin, an adipocyte-derived plasma protein, inhibits
 RT endothelial NF-kappaB signaling through a camp-dependent pathway.";
 RL Circulation 102:1296-1301(2000).
 RN [6]
 RP FUNCTION.
 RX MEDLINE=21372498; PubMed=11479627;
 RA Yamauchi T., Kamon J., Waki H., Terauchi Y., Kubota N., Hara K.,
 RA Mori Y., Ide T., Murakami K., Tsuboyama-Kasaoka N., Ezaki O.,
 RA Akanuma Y., Gavrilova O., Vinson C., Reitman M.L., Kagechika H.,
 RA Shudo K., Yoda M., Nakano Y., Tobe K., Nagai R., Kimura S., Tomita M.,
 RA Froguel P., Kadowaki T.;
 RT "The fat-derived hormone adiponectin reverses insulin resistance
 RT associated with both lipodystrophy and obesity.";
 RL Nat. Med. 7:941-946(2001).
 RN [7]
 RP VARIANT ADIPONECTIN DEFICIENCY CYS-112.
 RX MEDLINE=20378830; PubMed=10918532;
 RA Takahashi M., Arita Y., Yamagata K., Matsukawa Y., Okutomi K.,
 RA Horie M., Shimomura I., Hotta K., Kuriyama H., Kihara S., Nakamura T.,
 RA Yamashita S., Funahashi T., Matsuzawa Y.;
 RT "Genomic structure and mutations in adipose-specific gene,
 RT adiponectin.";
 RL Int. J. Obes. Relat. Metab. Disord. 24:861-868(2000).
 RN [8]
 RP VARIANTS ARG-84; MET-117; THR-164; SER-221 AND PRO-241.
 RX MEDLINE=21671103; PubMed=11812766;
 RA Hara K., Boutin P., Mori Y., Tobe K., Dina C., Yasuda K., Yamauchi T.,
 RA Otobe S., Okada T., Eto K., Kadowaki H., Hagura R., Akanuma Y.,
 RA Yazaki Y., Nagai R., Taniyama M., Matsubara K., Yoda M., Nakano Y.,
 RA Kimura S., Tomita M., Kimura S., Ito C., Froguel P., Kadowaki T.;
 RT "Genetic variation in the gene encoding adiponectin is associated with
 RT an increased risk of type 2 diabetes in the Japanese population.";
 RL Diabetes 51:536-540(2002).
 CC -1- FUNCTION: IMPORTANT NEGATIVE REGULATOR IN HEMATOPOIESIS AND IMMUNE
 CC SYSTEMS; MAY BE INVOLVED IN ENDING INFLAMMATORY RESPONSES THROUGH
 CC ITS INHIBITORY FUNCTIONS. INHIBITS ENDOTHELIAL NF-KAPPAB SIGNALING
 CC THROUGH A CAMP-DEPENDENT PATHWAY. INHIBITS TNF-ALPHA-INDUCED
 CC EXPRESSION OF ENDOTHELIAL ADHESION MOLECULES. INVOLVED IN THE
 CC CONTROL OF FAT METABOLISM AND INSULIN SENSITIVITY.
 CC -1- SUBUNIT: HOMODIGOMER (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Synthesized exclusively by adipocytes and
 CC secreted into plasma.
 CC -1- DISEASE: Defects in APOL1 are the cause of adiponectin deficiency
 CC [MIM:605441], resulting in very low concentration of plasma
 CC adiponectin. Decreased adiponectin plasma levels are associated
 CC with obesity insulin resistance, and diabetes type 2.
 CC -1- PHARMACEUTICAL: Adiponectin might be used in the treatment of
 CC diabetes type 2 and insulin resistance.
 CC -1- SIMILARITY: Contains 1 collagenous domain.
 CC -1- SIMILARITY: Contains 1 C1Q domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way


```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D45371; BAA08227.1; -.
DR EMBL; AB012165; BAA86716.1; -.
DR EMBL; AB012164; BAA86716.1; JOINED.
DR EMBL; AJ131460; CAB52413.1; -.
DR EMBL; AJ131461; CAB52413.1; JOINED.
DR PIR; JC4708; JC4708.
DR MIM; 605441; -.
DR GO; GO:0006091; P:energy pathways; TAS.
DR InterPro; IPR01073; Clq.
DR InterPro; IPR00087; Collagen.
DR Pfam; PF00386; Clq; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR ProDom; PD000007; Clq_helix; 1.
DR SMART; SM00110; Clq; 1.
DR PROSITE; PS01113; Clq; 1.
KW Hormone; Collagen; Signal; Repeat; Hydroxylation; Plasma;
KW Polymorphism; Disease mutation; Obesity; Diabetes mellitus.
FT SIGNAL 1 14
FT CHAIN 15 244
FT DOMAIN 42 107
FT DOMAIN 108 244
FT DISULFID 36 36
FT MOD_RES 44 44
FT MOD_RES 47 47
FT MOD_RES 53 53
FT MOD_RES 62 62
FT MOD_RES 71 71
FT MOD_RES 76 76
FT MOD_RES 86 86
FT MOD_RES 95 95
FT MOD_RES 104 104
FT VARIANT 84 84
FT VARIANT 112 112
FT VARIANT 117 117
FT VARIANT 164 164
FT VARIANT 221 221
FT VARIANT 241 241
FT SEQUENCE 244 AA; 26414 MW; 64D8C6C1204B1018 CRC64;
Query Match 11.7%; Score 192; DB 1; Length 244;
Best Local Similarity 31.3%; Pred. No. 1.1e-08;
Matches 47; Conservative 30; Mismatches 55; Indels 18; Gaps 6;
QY 5 PGPGSELRSASFSAARTPLEGTSSEMAVTFDKYVYVNIIGDFDVATGQFRCRVPGAYFFSF 64
DB 104 PGEGAYVYRSASFVGLTYVT-IPNMPIRFTKIFYNQONHYDGSTGKFHCNIFGLYYFAY 162
QY 65 TAGKAPH-----KSLSVMLVRNDEVQALAFDEQRRPGARRAASQASAMLQLDYGDVWLK 119
DB 163 -----HITVYMKDVKVSLEFK-KDKAMLFTYDQYENNVQD-ASGSVLLHLEVQDQVWLQ 214
QY 120 LHGAPH---YALGAPGATFSGYLVYADAD 145
DB 215 VYGEGERNGLYADNDNDSTFTGFLYHDTN 244
RESULT 8
CIRF_HUMAN STANDARD; PRT; 258 AA.
AC 075973;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

```

```

DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Clq-related factor precursor.
GN ClQRF OR CRF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99097006; PubMed=9878755;
RA Berube N.G., Swanson X.H., Bertram M.J., Kittle J.D., Didenko V.,
RA Baekin D.S., Smith J.R., Pereira-Smith O.M.;
RT "Cloning and characterization of CRF, a novel Clq-related factor,
RT expressed in areas of the brain involved in motor function.";
RL Brain Res. Mol. Brain Res. 63:233-240(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA Nickerson D.A.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loguelli N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAINSTEM.
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 Clq domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF095154; AAC64186.1; -.
DR EMBL; AF410771; AAK95248.1; -.
DR EMBL; BC008798; AAH08798.1; -.
DR GO; GO:0007626; P:locomotory behavior; NAS.
DR InterPro; IPR01073; Clq.
DR InterPro; IPR00087; Collagen.
DR Pfam; PF00386; Clq; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; Clq; 1.
DR PROSITE; PS01113; Clq; 1.
KW Collagen; Signal.
FT SIGNAL 1 16
FT CHAIN 17 258
FT DOMAIN 67 115
FT DOMAIN 123 258
FT SEQUENCE 258 AA; 26452 MW; 52C51CDF59CAE2BF CRC64;

```

```

Query Match          11.7%; Score 192; DB 1; Length 258;
Best Local Similarity 27.3%; Pred. No. 1.2e-08;
Matches 67; Conservative 26; Mismatches 88; Indels 64; Gaps 9

QY      94 RRRGARRAASQSAMLQLDYGDVTWLRHLGHAPHVAL-----GAPGATFSGYLVYADADADA 148
       |||||
       |   |   :   |   |   |   |   |   |   |   |   |   |   |   |   |
Db      38 RGPAG-----ARTDGDALSEGSAPPSTLVQGPGKPGRT-----GKP 78

QY      149 PARCPPAPPEPRSAFSAARTRSLVGSACPG-----PR----- 181
       |||||
       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
Db      79 GPPGPPPDGPPGPPGVPPGEKGEPKGPDPGLPAGCGSIAISTATYTIVRYAFYAGLKN 138

QY      182 ----HQPFAFDTEFFVNIGGFDAAGVFRCLPGAYFFSFLLGKLPR---KLSVKLMK 233
       ::|||
       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
Db      139 PHEGYEVLKFDDVNTNLGNNDYDASGKTCNIPGTYYFTYHV--LMRGDGTSMWADLCK 196

QY      234 NRDEVQAMITYDDGASRRREMOSQSVMLALRGDVAWVLLSHDHGYGAVSNHGKYTFESGF 293
       |:|:|
       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
Db      197 N-GQVRASAIAQDADQNYYDYASNVSILHLDAGDEVEFIKLDGKAHGNNSN--KYSTFSGF 253

QY      294 LVYPD 298
       ::|||
Db      254 IITYSD 258

RESULT 9
APM1_MOUSE STANDARD; PRT; 247 AA.
ID APM1_MOUSE O62400; Q9DC68;
AC O60994; O62400; Q9DC68;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Adiponectin precursor (30 kDa adipocyte complement-related protein)
DE (ACRP30) (Adipocyte specific protein AdipoQ).
GN APM1 OR ACRP30 OR ADIPOQ.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adipocyte; PubMed=7592907;
RX MEDLINE=96070757; PubMed=7592907;
RA Scherer P.E., Williams S., Fogliano M., Baldini G., Lodish H.F.;
RT "A novel serum protein similar to Ctg, produced exclusively in
RT adipocytes."
RL J. Biol. Chem. 270:26746-26749(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fibroblast; PubMed=8631877;
RX MEDLINE=96209999; PubMed=8631877;
RA Hu F., Liang P., Spiegelman B.M.;
RT "AdipoQ is a novel adipose-specific gene dysregulated in obesity.";
RL J. Biol. Chem. 271:10697-10703(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX PubMed=1162643; PubMed=1162643;
RA Das K., Lin Y., Widen E., Zhang Y., Scherer P.E.;
RT "Chromosomal localization, expression pattern, and promoter analysis
RT of the mouse gene encoding adipocyte-specific secretory protein
RT Acrp30."
RL Biochem. Biophys. Res. Commun. 280:1120-1129(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyoasa H., Kondo S., Yamanaoka I.,
RA Saito T., Okasaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

```

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Batsch G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bernaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald C., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsuki S.,
RA Hayashizaki Y.;
RT "functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [5]
RP FUNCTION.
RX MEDLINE=21372498; PubMed=11479627;
RA Yamauchi T., Kamon J., Waki H., Teranuchi Y., Kubota N., Hara K.,
RA Mori Y., Ide T., Murakami K., Tsuboyama-Kasaoka N., Ezaki O.,
RA Akanuma Y., Gavrilova O., Vinson C., Reitman M.L., Kagechika H.,
RA Shudo K., Yoda M., Nakano Y., Tobe K., Nagai R., Kimura S., Tomita M.,
RA Froguel P., Kadowaki T.;
RT "The fat-derived hormone adiponectin reverses insulin resistance
associated with both lipodystrophy and obesity.";
RL Nat. Med. 7:941-946(2001).
RN [6]
RP FUNCTION.
RX MEDLINE=21372499; PubMed=11479628;
RA Berg A.H., Combs T.P., Du X., Brownlee M., Scherer P.E.;
RT "The adipocyte-secreted protein Acrp30 enhances hepatic insulin
action.";
RL Nat. Med. 7:947-953(2001).
CC -1- FUNCTION: IMPORTANT NEGATIVE REGULATOR IN HEMATOPOIESIS AND IMMUNE
CC SYSTEMS; MAY BE INVOLVED IN ENDING INFLAMMATORY RESPONSES THROUGH
CC ITS INHIBITORY FUNCTIONS. INHIBITS ENDOTHELIAL NF-KAPPAB SIGNALING
CC THROUGH A CAMP-DEPENDENT PATHWAY. INHIBITS TNF-ALPHA-INDUCED
CC EXPRESSION OF ENDOTHELIAL ADHESION MOLECULES. INVOLVED IN THE
CC CONTROL OF FAT METABOLISM AND INSULIN SENSITIVITY.
CC -1- SUBUNIT: Homooligomer.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Synthesized exclusively by adipocytes and
CC secreted into plasma.
CC -1- INDUCTION: DURING HORMONE-INDUCED ADIPOSE DIFFERENTIATION AND
CC ACTIVATED BY INSULIN.
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C1q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U37222; AAA80543.1; -;
DR EMBL; U49915; AAB06706.1; -;
DR EMBL; AF304466; AAK13417.1; -;
DR EMBL; AK003138; BAB22597.1; -;
DR PDB; 1C28; 07-SEP-99.
DR MGD; MGI:106675; Acrp30.
DR GO; GO:0005576; C:extracellular; IDA.
DR GO; GO:0005515; F:protein binding activity; IPI.
DR GO; GO:0006635; P:fatty acid beta-oxidation; IMP.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR ProDom; PD000007; C1q_helix; 1.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
DR Hormone; Collagen; Signal; Repeat; Hydroxylation; Plasma;
KW

| KW | Polymorphism; 3D-structure. | POTENTIAL. |
|----|--|--------------------------------|
| FT | SIGNAL 1 17 | ADIPONECTIN. |
| FT | CHAIN 18 247 | COLLAGEN-LIKE. |
| FT | DOMAIN 45 110 | C10. |
| FT | DOMAIN 111 247 | INTERCHAIN (BY SIMILARITY). |
| FT | DISULFID 39 39 | HYDROXYLATION (BY SIMILARITY). |
| FT | MOD_RES 47 47 | HYDROXYLATION (BY SIMILARITY). |
| FT | MOD_RES 50 50 | HYDROXYLATION (BY SIMILARITY). |
| FT | MOD_RES 56 56 | HYDROXYLATION (BY SIMILARITY). |
| FT | MOD_RES 65 65 | HYDROXYLATION (BY SIMILARITY). |
| FT | MOD_RES 79 79 | HYDROXYLATION (BY SIMILARITY). |
| FT | MOD_RES 98 98 | HYDROXYLATION (BY SIMILARITY). |
| FT | MOD_RES 107 107 | HYDROXYLATION (BY SIMILARITY). |
| FT | VARIANT 113 113 | M -> V. |
| FT | CONFLICT 50 50 | P -> S (IN REF. 2). |
| FT | CONFLICT 74 74 | A -> S (IN REF. 2). |
| FT | CONFLICT 117 117 | A -> G (IN REF. 2). |
| FT | CONFLICT 148 148 | G -> N (IN REF. 2). |
| FT | CONFLICT 243 243 | Y -> F (IN REF. 2). |
| SO | SEQUENCE 247 AA; 26841 MW; 137B687D873988C4 CRC64; | |

| | | | | |
|--------------------------|-------|--------------------|-----------|-------------|
| Query Match | 11.5% | Score 188; | DB 1; | Length 247; |
| Best Local Similarity | 33.5% | Pred. No. 2.3e-08; | | |
| Matches 53; Conservative | 24; | Mismatches 73; | Indels 8; | Gaps 5; |

Qy 146 ADAPARGPPAPP----EPRSAFSAARTSLVGS DAGGPRHQPLAFDTEFVNIGDFDA 201
| : | | | : | : | : | : | :
Db 91 AEGP-RGFPGTPGRKGEPEGAAYMYRSASFVGLSETRVTVPNPPIRFTKIIFYNQNHYDGS 149

```

Qy      202 AGVFRCLPGAYFFSFTLGLPKRTL SVKLMKNRDEVQAMIYDDGASRRREMOSQVMLA 261
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      150 TGKEYCNIPGLYYFSYHI -TVYMKDVKYSLEK -KDKAVLFTYDQYQEKNVDAQSGSVLLH 207

```

```
QY      262 LRRGDAVWLISH-DHDCYGAYSNHGKYITFSGFLVYPD 298
        ||||| : | | | | : : | | | | : |
Db      208 LEVGDDVWLQVYGGDGHNGLIADNVNDSTFTGFLIYHD 245
```

| RESULT 10 | |
|---------------|------------------------|
| GLIC_MOUSE | |
| ID_GLIC_MOUSE | STANDARD; PRT: 255 AA. |

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glucocorticoid precursor (C19-like protein).
GN C19L.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
OX NCBI TaxID=10090;

RP SEQUENCE FROM N.A.
RX MEDLINE=20428709; PubMed=10862616;
RA Koide T., Abo A., Yoriizu T., Nagata K.;
RT "Conformational requirements of collagenous peptides for recognition
RL by the chaperone protein HSP47.";
J. Biol. Chem. 275:27957-27963 (2000).

CC -1- TISSUE SPECIFICITY: EXPRESSED IN GLIAL CELLS.
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C1Q domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).

DR EMBL; AB044560; BAB15806.1; -.
DR MGD; MG1:2387350; C1q1.
DR GO; GO:0005515; F:protein binding activity; IPI.

| | |
|----|--|
| DR | InterPro; IPR001073; Clq. |
| DR | InterPro; IPR000087; Collagen. |
| DR | Pfam; PF00386; Clq; 1. |
| DR | Pfam; PF01391; Collagen; 1. |
| DR | PRINTS; PR00007; COMPLEMENTC1Q. |
| DR | SMART; SM00110; ClQ; 1. |
| DR | PROSITE; PS01113; Clq; 1. |
| KW | Collagen; Signal. |
| FT | SIGNAL 1 20 POTENTIAL. |
| FT | CHAIN 21 255 GLIACOLIN. |
| FT | DOMAIN 61 111 COLLAGEN-LIKE. |
| FT | DOMAIN 120 255 ClQ. |
| SQ | SEQUENCE 255 AA; 26687 MW; 529FBAF4B2191BC1 CRC64; |

| | | | | | |
|-----------------------|------------------|--------------------|------------|-------------|--|
| Query Match | 11.4% | Score 186.5; | DB 1; | Length 255; | |
| Best Local Similarity | 26.3%; | Pred. No. 3.2e-08; | | | |
| Matches 57; | Conservative 23; | Mismatches 86; | Indels 51; | Gaps 6; | |

OY 120 LHGAPHYALGAPGATGSGYLTVADADADAPARGPPAPEPRSAFSAARTSLVGS D A P G 179
 : | : | ||| : |
 Db 52 MQLPTFIQGPKGE-----AGRPCKAGPRGPPEPVGPPGSEKCEPGRQGLPG 103

```

QY      180 PRHQ-----LAFDTEFVNI GGFEDAAAGVF 205
          | | | | | | | | | |
Db      104 PRGAPGLNAGATSAATYSTVPRKIAFYAGLKRQHEGYEVLKFDVDVTNLGNHYDPTTGKF 163

```

[illegible]

```

OY      262 LRRGDVAVLLSHDHDGYGAYSNHGKYYITFSGFLVYPD 298
      | | | | | : | | | | | : | | | | |
DB      221 LEPGDEVYIKLDGSKAHG--GNNNNKYSTFSGFIIVAD 255

```

| | |
|---------------|-------------|
| RESULT 11 | |
| HP25 TAMSI | |
| ID HP25 TAMSI | STANDARD; |
| | PRT; 215 AA |

DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, last sequence update)
DT 28-FEB-2003 (Rel. 41, last annotation update)
DE Hibernator-associated plasma protein HP-25 precursor (Hibernator-specific blood complex, 25 kDa subunit).

OS *Tamias sibiricus* (Siberian chipmunk) (Asian chipmunk).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Scurinae;
 OC *Tamias*.

```

OX  NCBI_TaxID=64680;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=liver;

```

RX MEDLINE=93180798; PubMed=8441393;
RA Takamatsu N., Ohba K., Kondo J., Kondo N., Shiba T.;
RT "Hibernation-associated gene regulation of plasma proteins with a
RT collagen-like domain in mammalian hibernators.";
RL Mol. Cell. Biol. 13:1516-1521(1993).

RP SEQUENCE OF 29-62; 84-130; 172-183; 187-192 AND 201-215.
RC TISSUE=Plasma;
RX MEDLINE=92112696; PubMed=1730610;

RA Kondo N', Kondo J';
RT "Identification of novel blood proteins specific for mammalian
RT hibernation."; *J. Biol. Chem.* 267:473-478(1992).
RL
CC -1- FUNCTION: PLASMA PROTEINS HP-20, HP-25, HP-27 AND HP-55 FORM A
CC 140 Kda COMPLEX VIA DISULFIDE BONDS IN THE PLASMA AND ARE
CC HIBERNATION SPECIFIC.
CC -1- SUBCELLULAR LOCATION: Extracellular.

CC -1- DEVELOPMENTAL STAGE: THE PROTEIN COMPLEX DISAPPEARS FROM THE
CC PLASMA AT ONSET OF HIBERNATION AND REAPPEARS AS HIBERNATION
CC CEASES.

CC -1- SIMILARITY: Contains 1 collagenous domain.
 CC -1- SIMILARITY: Contains 1 C1Q domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D12975; BAA02352.1; -
 DR PIR; B48150; B48150.
 DR InterPro; IPR001073; C1q.
 DR InterPro; IPR000087; Collagen.
 DR Pfam; PF00386; C1q; 1.
 DR Pfam; PF01391; Collagen; 1.
 DR PRINTS; PR00007; COMPLEMENTC1Q.
 DR SMART; SM00110; C1Q; 1.
 DR PROSITE; PS01113; C1Q; 1.
 KW Signal; Collagen; Glycoprotein; Plasma; Multigene family.
 FT SIGNAL 1 28
 FT CHAIN 29 215 HIBERNATION-ASSOCIATED PLASMA PROTEIN
 FT HP-25.
 FT DOMAIN 40 81 COLLAGEN-LIKE.
 FT DOMAIN 83 215 C1Q.
 FT CARBOHYD 167 167 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 215 AA; 22664 MW; AFE03206917EA530 CRC64;
 Query Match 11.2%; Score 184; DB 1; Length 215;
 Best Local Similarity 30.4%; Pred. No. 4.2e-08;
 Matches 62; Conservative 22; Mismatches 52; Indels 68; Gaps 11;
 QY 122 GAPHYALGAPGATFSGYLVYADADADAPARGPAPP-----EPRS 161
 DB 49 GIPGFP-GAPGAL-----GPPGPPGVGIPGPQGPDPVEKCSSRPKS 90
 QY 162 AFSAAARTSLVGS DAGPGRHOPLAFDTFVNIGD FDAAGVFRCLPGAYFFSFTLGK 221
 DB 91 AFAVKL-----SERPPEP-FQPIVFKEALYNGEHFMATGEFSCVLPGVNFGFDI-R 142
 QY 222 LPRKTL SVKLMKNRDEVA MIYDDGASRRREMOSQ-----SVMLALRGDAVWL LSH 273
 DB 143 LFQSSVKIRLM--RDGIQV-----REKEAQANDSYKHAMGSVIMALGKDKVWLBSK 192
 QY 274 DHDGYGAYSNG-KYITFSGFLVY 296
 DB 193 LK--GTESEKGI THIVFFGYLLY 213
 RESULT 12
 C1QB HUMAN STANDARD; PRT; 251 AA.
 ID C1QB_HUMAN
 AC P02746; Q96H17;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Complement C1q subcomponent, B chain precursor.
 GN C1QB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=86076906; PubMed=3000358;
 RX Reid K.B.M.;
 RA "Molecular cloning and characterization of the complementary DNA and
 RT gene coding for the B-chain of subcomponent C1q of the human
 RT complement system."
 RL Biochem. J. 231:729-735(1985).
 RN [2]
 RP SEQUENCE FROM N.A.

RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE OF 26-133.
 RX MEDLINE=80020137; PubMed=486087;
 RA Reid K.B.M.;
 RT "Complete amino acid sequences of the three collagen-like regions
 RT present in subcomponent C1q of the first component of human
 RT complement."
 RL Biochem. J. 179:367-371(1979).
 RN [4]
 RP SEQUENCE OF 26-193.
 RX MEDLINE=79041552; PubMed=708376;
 RA Reid K.B.M., Thompson E.O.P.;
 RT "Amino acid sequence of the N-terminal 108 amino acid residues of the
 RT B chain of subcomponent C1q of the first component of human
 RT complement."
 RL Biochem. J. 173:863-868(1978).
 RN [5]
 RP SEQUENCE OF 134-251.
 RX MEDLINE=82283890; PubMed=6981411;
 RA Reid K.B.M., Gagnon J., Frampton J.;
 RT "Completion of the amino acid sequences of the A and B chains of
 RT subcomponent C1q of the first component of human complement."
 RL Biochem. J. 203:559-569(1982).
 RN [6]
 RP SEQUENCE OF 224-251 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=85038855; PubMed=6208566;
 RA Reid K.B.M., Bentley D.R., Wood K.J.;
 RT "Cloning and characterization of the complementary DNA for the B
 RT chain of normal human serum C1q."
 RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 306:345-354(1984).
 RN [7]
 RP REVIEW OF C1Q DEFICIENCY.
 RX MEDLINE=98450587; PubMed=9777412;
 RA Petry F.;
 RT "Molecular basis of hereditary C1q deficiency."
 RL Immunobiology 199:286-294(1998).
 CC -1- FUNCTION: C1Q ASSOCIATES WITH THE PROENZYMES C1R AND C1S TO YIELD
 CC C1, THE FIRST COMPONENT OF THE SERUM COMPLEMENT SYSTEM. THE
 CC COLLAGEN-LIKE REGIONS OF C1Q INTERACT WITH THE CA(2+)-DEPENDENT
 CC C1R(2)C1S(2) PROENZYME COMPLEX, AND EFFICIENT ACTIVATION OF C1
 CC TAKES PLACE ON INTERACTION OF THE GLOBULAR HEADS OF C1Q WITH THE
 CC FC REGIONS OF IGG OR IGM ANTIBODY PRESENT IN IMMUNE COMPLEXES.
 CC -1- SUBUNIT: C1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q, R
 CC AND S IN THE MOLAR RATION OF 1:2:2. C1Q SUBCOMPONENT IS COMPOSED
 CC OF NINE SUBUNITS, SIX OF WHICH ARE DISULFIDE-LINKED DIMERS OF THE
 CC A AND B CHAINS, AND THREE OF WHICH ARE DISULFIDE-LINKED DIMERS OF
 CC THE C CHAIN.
 CC -1- PTM: O-linked glycans consist of Glc-Gal disaccharides bound to
 CC the oxygen atom of post-translationally added hydroxyl groups.
 CC -1- DISEASE: Defects in C1QB are a cause of C1q deficiency

[MIM:120570]. It is a rare genetic disorder which is associated with recurrent infections and a high prevalence of lupus erythematosus-like symptoms. It is characterized by a loss of activation of the complement classical pathway.

-1- SIMILARITY: Contains 1 collagenous domain.

-1- SIMILARITY: Contains 1 C1q domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X03084; CAA26880.1; -. ALT_INIT.
EMBL; BC008983; AAH08983.1; ALT_INIT.
EMBL; M36278; AAC41692.1; -.
Gene; HGNC:1242; C1QB.
MIM; 120570; -.
GO; GO:0005602; C:complement component C1q complex; TAS.
GO; GO:0003811; F:complement activity activity; TAS.
InterPro; IPR001073; C1q.
InterPro; IPR000087; Collagen.
Pfam; PF00386; C1q; 1.
Pfam; PF01391; Collagen; 1.
PRINTS; PR00007; COMPLEMENTC1Q.
SMART; SM00110; C1Q; 1.
PROSITE; PS01113; C1Q; 1.
KW Complement pathway; Plasma; Hydroxylation; Glycoprotein; Collagen; Repeat; Signal; Disease mutation; Pyrrolidone carboxylic acid.
FT SIGNAL 1 25
FT CHAIN 26 251 COMPLEMENT C1Q SUBCOMPONENT, B CHAIN.
FT DOMAIN 29 112 COLLAGEN-LIKE.
FT DOMAIN 113 251 C1Q.
FT MOD_RES 26 26 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 29 29 INTERCHAIN (WITH C-26 IN CHAIN A).
FT MOD_RES 33 33 HYDROXYLATION.
FT MOD_RES 36 36 HYDROXYLATION.
FT MOD_RES 39 39 HYDROXYLATION.
FT MOD_RES 42 42 HYDROXYLATION.
FT MOD_RES 51 51 HYDROXYLATION.
FT MOD_RES 54 54 HYDROXYLATION.
FT MOD_RES 57 57 HYDROXYLATION.
FT CARBOHYD 57 57 O-LINKED (GAL. . .).
FT MOD_RES 60 60 HYDROXYLATION.
FT CARBOHYD 60 60 O-LINKED (GAL. . .).
FT MOD_RES 63 63 HYDROXYLATION.
FT MOD_RES 75 75 HYDROXYLATION.
FT MOD_RES 81 81 HYDROXYLATION.
FT MOD_RES 84 84 HYDROXYLATION.
FT MOD_RES 90 90 HYDROXYLATION.
FT MOD_RES 96 96 HYDROXYLATION.
FT CARBOHYD 96 96 O-LINKED (GAL. . .).
FT MOD_RES 99 99 HYDROXYLATION.
FT MOD_RES 102 102 HYDROXYLATION.
FT MOD_RES 105 105 HYDROXYLATION.
FT MOD_RES 108 108 HYDROXYLATION.
FT CARBOHYD 108 108 O-LINKED (GAL. . .).
FT VARIANT 40 40 G -> D (in C1Q deficiency).
FT 26 26 /FTID=VAR 008541.
FT CONFLICT 83 83 Q -> E (IN REF. 3).
FT CONFLICT 98 98 N -> D (IN REF. 3).
SQ SEQUENCE 251 AA; 26459 MW; 78C5752E267A0EF7 CRC64;
Query Match 11.1%; Score 182; DB 1; Length 251;
Best Local Similarity 29.1%; Pred. No. 7.3e-08;
Matches 48; Conservative 32; Mismatches 61; Indels 24; Gaps 5;

QY 149 PARGPAPPEP-----RSAFSAARTSLVSGSDAGPGRHOPLAFDTEFNIG 195
DB 95 PKGGPAGAPGPKGSGDYKATOKIAFSATRTINV-----PLRRDQTRPDHVTNNM 148

QY 196 GDFDAAGVRCRLPGAYFFSFTLGKLPKTLVVKLMKNRDEVQAM1-YDGSARRREMQ 254
DB 149 NNYEPRSGKFTCKVPLGYFTYTHASS--RGNLCVNLMRGRERAQKVTFCDYAYNTFQVT 206

QY 255 SOSVMLALRGDAVWLSHDHDGYGAYSNHGKYYTFSGFLVYPPD 299
DB 207 TGGMVLTKEGENVFLQATDKN--SLIGMEGANSIFSGFLLFPDM 249

RESULT 13
CQT7 HUMAN STANDARD; PRT; 289 AA.
ID CQT7_HUMAN
AC O9BXJ2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement-clq tumor necrosis factor-related protein 7 precursor.
GN C1QTNF7 OR CTRP7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Piddington C.S., Sheppard P.O., Bishop P., Lasser G.W.,
RT "Homo sapiens complement-clq tumor necrosis factor-related protein.";
RL Submitted (DEC-2000) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; Pubmed=12477932;
RA Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Keltman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C1q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----

EMBL; AF329839; AAK17963.1; -.
EMBL; BC022187; AAH22187.1; -.
Gene; HGNC:14342; C1QTNF7.
InterPro; IPR001073; C1q.
InterPro; IPR000087; Collagen.
Pfam; PF00386; C1q; 1.
Pfam; PF01391; Collagen; 2.
PRINTS; PR00007; COMPLEMENTC1Q.
SMART; SM00110; C1Q; 1.
PROSITE; PS01113; C1Q; 1.

```
KW Collagen; signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 289 COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT DOMAIN 38 139 RELATED PROTEIN 7.
FT DOMAIN 141 276 COLLAGEN-LIKE.
SQ SEQUENCE 289 AA; 30683 MW; A61609FF6BD26946 CRC64;

Query Match 11.0%; Score 181; DB 1; Length 289;
Best Local Similarity 33.7%; Pred. No. 1e-07;
Matches 57; Conservative 21; Mismatches 59; Indels 32; Gaps 8;

QY 1 LGPTPGP-----GSSELRSAFSAARTT--PLEGTSMAVTFDKVYN 40
DB 118 IGP-PGPKGDRGEQDPLPGVCRGSIVLKSAFSGITTSYPEE--RLPIIFNKVLFN 173
QY 41 IGDGFDVATGQRCRVPGAYFFSFAGKAPHSLSVMLVRNRDEVQALAFDEQRRPGARR 100
DB 174 EGEHYNPATGKFCIFPGIYFYSYDITLA-NKHLAIGLVHN-GQYRIKTFD--ANTGNHD 229
QY 101 AASQSAMLQLDYGTWVLRHLGAPHYAL---GAPGATFSGYLVYADAD 145
DB 230 VASGSTVIYVLPDEVEVLEIFFTDQNGLFSDPGWADSLFSGFLLYVDTD 278

RESULT 14
CA18 RABIT STANDARD; PRT; 744 AA.
ID CA18 RABIT STANDARD; PRT; 744 AA.
AC P14282;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Collagen alpha 1(VIII) chain precursor (Endothelial collagen).
GN COL8A1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89380199; PubMed=2476437;
RA Yamaguchi N., Benya P.D., van der Rest M., Ninomiya Y.;
RT "The cloning and sequencing of alpha 1(VIII) collagen cDNAs
RT demonstrate that type VIII collagen is a short chain collagen and
RT contains triple-helical and carboxyl-terminal non-triple-helical
RT domains similar to those of type X collagen.";
RL J. Biol. Chem. 264:16022-16029(1989).
CC -1- FUNCTION: MAJOR COMPONENT OF THE DESCMET'S MEMBRANE (BASEMENT
CC MEMBRANE) OF CORNEAL ENDOTHELIAL CELLS.
CC -1- SUBUNIT: MAY FORM HOMOTRIMERS, OR HETEROTRIMERS IN ASSOCIATION
CC WITH ALPHA 2(VIII) TYPE COLLAGENS.
CC -1- PTM: Prolines at the third position of the tripeptide repeating
CC unit (G-X-Y) are hydroxylated in some or all of the chains.
CC -1- MISCELLANEOUS: 4 consecutive G-P-P tripeptides are present at the
CC C-terminus of the triple-helical region. These may provide high
CC thermal stability of this region.
CC -1- SIMILARITY: STRONG, TO ALPHA 2 TYPES VIII AND X COLLAGENS.
CC -1- SIMILARITY: Contains 1 C1Q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J05042; AAA31204.1; -
CC PIR; A34246; A34246.
CC InterPro; IPR001073; C1q.
CC InterPro; IPR000087; Collagen.
CC Pfam; PF00386; C1q; 1.
CC Pfam; PF01391; Collagen; 8.
```

```
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Glycoprotein; Cell adhesion; Collagen; signal.
FT SIGNAL 1 20
FT CHAIN 21 744 COLLAGEN ALPHA 1(VIII) CHAIN.
FT DOMAIN 29 117 NONHELICAL REGION (NC2).
FT DOMAIN 118 571 TRIPLE-HELICAL REGION.
FT DOMAIN 572 744 NONHELICAL REGION (NC1).
FT DOMAIN 609 744 C1Q.
SQ SEQUENCE 744 AA; 73358 MW; 2A8CEF1EF8274E99 CRC64;

Query Match 10.6%; Score 173; DB 1; Length 744;
Best Local Similarity 25.3%; Pred. No. 1.4e-06;
Matches 56; Conservative 27; Mismatches 78; Indels 60; Gaps 8;

QY 120 LHGAPHYALGAPGATFSGYLVYADADADAPA-RGPPAPPEPRs----- 161
DB 540 LHGPP-----GKPGA-----LGPGQGLPGPPGPPGPPAPVMPPTPAPQGEYLPDMG 589
QY 162 -----AFSAARTSLVSGDAGG-----PRHQLAFDTEFVNIGG 196
DB 590 LGIDGVKTPHAYAAKK-----GKNGPAYEMPAFTALPTAPPVGAPIKFDRLLYNGRQ 644
QY 197 DFDAAGVFRCLPGAYFFSFTLGKLPKRTLsvLmKNRDEVQAMTYDDGASRRREMOSQ 256
DB 645 NYNPQTGTFCEVPGVYFYAYHV-HCKGNVWVALLFKNNEPVM-YTYDEVKKGLDQASG 702
QY 257 SVMALRRGDVAVLLSHDHGCGAYSNHGKXITFSGFLVYP 297
DB 703 SAVLLLRPDGRVFLQMPSEQAAGLYAGQYVHSSFSGYLLYP 743

RESULT 15
COLE LEPPA STANDARD; PRT; 419 AA.
ID COLE LEPPA STANDARD; PRT; 419 AA.
AC P98085; Q91080;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Inner ear-specific collagen precursor (Saccular collagen).
OS Lepomis macrochirus (Bluegill).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Centrarchidae; Lepomis.
OX NCBI_TaxID=13106;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95167486; PubMed=7863331;
RA Davis J.G., Oberholtzer J.C., Burns F.R., Greene M.I.;
RT "Molecular cloning and characterization of an inner ear-specific
RT structural protein.";
RL Science 267:1031-1034(1995).
RN [2]
RP CONCEPTUAL TRANSLATION.
RA Gibson T.;
RL Submitted (MAR-1995) to the SWISS-PROT data bank.
CC -1- FUNCTION: FORMS A MICROSTRUCTURAL MATRIX WITHIN THE OTOLITHIC
CC MEMBRANE (PROBABLY).
CC -1- TISSUE SPECIFICITY: SPECIALIZED SECRETORY SUPPORTING CELLS AT THE
CC OUTER PERIMETER OF THE SACCULAR EPITHELIUM.
CC -1- SIMILARITY: Contains 1 C1Q domain.
CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A FRAMESHIFT HAD TO BE
CC INTRODUCED FOR POSITIONS 391-419 SO AS TO MAXIMIZE THE SIMILARITY
CC WITH OTHER SHORT-CHAIN COLLAGENS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

| | |
|----|--|
| DR | EMBL; U17431; AAA69978.1; ALT_FRAME. |
| DR | InterPro; IPR001073; C1q. |
| DR | InterPro; IPR000087; Collagen. |
| DR | Pfam; PF00386; C1q; 1. |
| DR | Pfam; PF01391; Collagen; 3. |
| DR | PRINTS; PR00007; COMPLEMENTC1Q. |
| DR | ProDom; PD000007; C1q_helix; 2. |
| DR | SMART; SM00110; C1Q; 1. |
| DR | PROSITE; PS01113; C1Q; 1. |
| KW | Extracellular matrix; Repeat; Collagen; Signal. |
| FT | SIGNAL 1 19 POTENTIAL. |
| FT | CHAIN 20 419 INNER EAR-SPECIFIC COLLAGEN. |
| FT | DOMAIN 20 57 NONHELICAL REGION (NC2). |
| FT | DOMAIN 58 274 TRIPLE-HELICAL REGION (COL1). |
| FT | DOMAIN 275 419 NONHELICAL REGION (NC1). |
| FT | DOMAIN 272 419 C1Q. |
| FT | CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL). |
| SO | SEQUENCE 419 AA; 43634 MW; 570CDB9675FC0F39 CRC64; |

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:10 ; Search time 43.1078 Seconds
(without alignments)
1873.686 Million cell updates/sec

Title: US-10-085-167-2_COPY_17_329

Perfect score: 1639

Sequence: 1 LGPTPGGSSSELSAFAFSAR.....LVYPDLAPAPGIGASELL 313

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SP_TREMBL_23:*
- 2: sp_archaea:*
- 3: sp_bacteria:*
- 4: sp_fungi:*
- 5: sp_human:*
- 6: sp_invertebrate:*
- 7: sp_mammal:*
- 8: sp_mhc:*
- 9: sp_organelle:*
- 10: sp_phage:*
- 11: sp_plant:*
- 12: sp_rodent:*
- 13: sp_virus:*
- 14: sp_vertebrate:*
- 15: sp_unclassified:*
- 16: sp_virus:*
- 17: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|--------|--------------------|
| 1 | 1631 | 99.5 | 329 | 4 | Q8IV25 | Q8IV25 homo sapien |
| 2 | 1514 | 92.4 | 326 | 11 | Q8R066 | Q8R066 mus musculu |
| 3 | 1003.5 | 61.2 | 205 | 11 | Q9D0W2 | Q9D0W2 mus musculu |
| 4 | 997.5 | 60.9 | 205 | 11 | Q9DCB6 | Q9DCB6 mus musculu |
| 5 | 230 | 14.0 | 158 | 4 | Q9H667 | Q9H667 homo sapien |
| 6 | 230 | 14.0 | 1077 | 4 | Q8TE71 | Q8TE71 homo sapien |
| 7 | 226 | 13.8 | 158 | 11 | Q8K110 | Q8K110 mus musculu |
| 8 | 220.5 | 13.5 | 182 | 11 | Q8R1P2 | Q8R1P2 mus musculu |
| 9 | 220.5 | 13.5 | 281 | 11 | Q9QXP7 | Q9QXP7 mus musculu |
| 10 | 213 | 13.0 | 243 | 4 | Q8N6P2 | Q8N6P2 homo sapien |
| 11 | 205 | 12.5 | 243 | 11 | Q8R002 | Q8R002 mus musculu |
| 12 | 202 | 12.3 | 243 | 11 | Q8K479 | Q8K479 mus musculu |
| 13 | 200 | 12.2 | 294 | 11 | Q9D8U4 | Q9D8U4 mus musculu |
| 14 | 193 | 11.8 | 243 | 6 | Q95JD7 | Q95JD7 macaca mula |
| 15 | 189 | 11.5 | 247 | 11 | Q8BRW2 | Q8BRW2 mus musculu |
| 16 | 186 | 11.3 | 287 | 11 | Q8CFR0 | Q8CFR0 mus musculu |

| | | | | | | |
|----|-------|------|------|----|--------|---------------------|
| 17 | 185 | 11.3 | 240 | 6 | Q95MQ4 | Q95mq4 bos taurus |
| 18 | 181.5 | 11.1 | 244 | 11 | Q8K3R4 | Q8K3r4 rattus norv |
| 19 | 180.5 | 11.0 | 264 | 11 | Q8BKRO | Q8Bkr0 mus musculu |
| 20 | 179.5 | 11.0 | 120 | 11 | Q8R1Z2 | Q8R1z2 mus musculu |
| 21 | 177.5 | 10.8 | 312 | 11 | Q8CHX9 | Q8chx9 mus musculu |
| 22 | 168.5 | 10.3 | 289 | 11 | Q8BVD7 | Q8bvd7 mus musculu |
| 23 | 167.5 | 10.2 | 744 | 11 | Q921S8 | Q921s8 mus musculu |
| 24 | 167.5 | 10.2 | 744 | 11 | Q8BGL6 | Q8bgl6 mus musculu |
| 25 | 166 | 10.1 | 675 | 6 | Q9N178 | Q9n178 sus scrofa |
| 26 | 165.5 | 10.1 | 194 | 6 | Q95J95 | Q95j95 canis fami1 |
| 27 | 164 | 10.0 | 295 | 11 | Q9Z1K4 | Q9z1k4 rattus norv |
| 28 | 163 | 9.9 | 224 | 4 | Q8IUK8 | Q8iuk8 homo sapien |
| 29 | 163 | 9.9 | 224 | 11 | Q8BGU2 | Q8bgu2 mus musculu |
| 30 | 162.5 | 9.9 | 246 | 11 | Q9ES30 | Q9es30 mus musculu |
| 31 | 161 | 9.8 | 173 | 6 | Q62789 | Q62789 sus scrofa |
| 32 | 159 | 9.7 | 245 | 11 | Q9DCM6 | Q9dcm6 mus musculu |
| 33 | 155 | 9.5 | 196 | 11 | Q920N0 | Q920n0 tamias sibi |
| 34 | 152 | 9.3 | 197 | 11 | Q9JHG0 | Q9jhg0 mus musculu |
| 35 | 147.5 | 9.0 | 333 | 4 | Q8IUT4 | Q8iut4 homo sapien |
| 36 | 145.5 | 8.9 | 705 | 4 | Q8TEJ5 | Q8tej5 homo sapien |
| 37 | 141.5 | 8.6 | 1017 | 11 | Q99K41 | Q99k41 mus musculu |
| 38 | 140 | 8.5 | 198 | 11 | Q8BMF0 | Q8bmf0 mus musculu |
| 39 | 139 | 8.5 | 198 | 11 | Q8BME9 | Q8bme9 mus musculu |
| 40 | 135 | 8.2 | 195 | 11 | Q8BZS3 | Q8bz s3 mus musculu |
| 41 | 134.5 | 8.2 | 213 | 5 | P83425 | P83425 mytilus edu |
| 42 | 131.5 | 8.0 | 347 | 4 | Q96IH6 | Q96ih6 homo sapien |
| 43 | 131.5 | 8.0 | 583 | 4 | Q96G58 | Q96g58 homo sapien |
| 44 | 131.5 | 8.0 | 992 | 4 | Q9UG76 | Q9ug76 homo sapien |
| 45 | 131.5 | 8.0 | 1016 | 4 | Q9Y6C2 | Q9y6c2 homo sapien |

ALIGNMENTS

| RESULT 1 | ID | Q8IV25 | PRELIMINARY; | PRT; | 329 AA. |
|----------|---|--------------------------------------|--------------|------|---------|
| AC | Q8IV25 | 01-MAR-2003 (TREMBLrel. 23, Created) | | | |
| DT | 01-MAR-2003 (TREMBLrel. 23, Last sequence update) | | | | |
| DT | 01-MAR-2003 (TREMBLrel. 23, Last annotation update) | | | | |
| DE | Similar to C1q and tumor necrosis factor related protein 4. | | | | |
| OS | Homo sapiens (Human). | | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | |
| OX | NCBI_TaxID=9606; | | | | |
| RN | [1] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RC | TISSUE=Brain; | | | | |
| RA | Strausberg R.; | | | | |
| RL | Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases. | | | | |
| DR | EMBL; BC035628; AAH35628.1; .. | | | | |
| SQ | SEQUENCE 329 AA; 35256 MW; 16064DA8182A6732 CRC64; | | | | |

| Query Match | Best Local Similarity | 99.7%; | Score 1631; | DB 4; | Length 329; |
|-------------|-----------------------|--|-------------|------------|-------------|
| Matches | 312; | Conservative | 0; | Mismatches | 1; |
| Indels | 0; | Gaps | 0; | | |
| QY | 1 | LGPTPGGSSSELSAFAFSARTTPLEGTSEMAVTFDKVYVNIIGDFVATGQFRCRVP | 60 | | |
| Db | 17 | LGPTPGGSSSELSAFAFSARTTPLEGTSEMAVTFDKVYVNIIGDFVATGQFRCRVP | 76 | | |
| QY | 61 | FFSFTAGKAPKSLSLVLRNRDEVQALAFDEQRRPGARRAASQAMQLDYGDTWLR | 120 | | |
| Db | 77 | FFSFTAGKAPKSLSLVLRNRDEVQALAFDEQRRPGARRAASQAMQLDYGDTWLR | 136 | | |
| QY | 121 | HGAPHYALGAPGATFSGYLVYADADADAPARGPAPPEPRSAFSAARTSLVGS | 180 | | |
| Db | 137 | HGAPQYALGAPGATFSGYLVYADADADAPARGPAPPEPRSAFSAARTSLVGS | 196 | | |
| QY | 181 | RHQPPLAFDTEFVNIIGDFAAAGVFRCPGAYFFSFTLGKLPKTLVYKLMKNRDE | 240 | | |
| Db | 197 | RHQPPLAFDTEFVNIIGDFAAAGVFRCPGAYFFSFTLGKLPKTLVYKLMKNRDE | 256 | | |

STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guernicich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlschki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK002948; BAB22473.1; -;
DR MGI; MGI:1914695; 0710001E10Rik.
DR InterPro; IPR001073; Clq.
DR Pfam; PF00386; Clq; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
SQ SEQUENCE 205 AA; 22215 MW; 56AD37793C437300 CRC64;

Query Match 60.9%; Score 997.5; DB 11; Length 205;
Best Local Similarity 93.2%; Pred. No. 7.4e-78;
Matches 192; Conservative 1; Mismatches 12; Indels 1; Gaps 1;
QY 107 MLQLDYGDTWLRHLGAPHYALGAPGATFSGYLYVADADADAPARGPPAPPEPRSAFSA 166
DB 1 MLQLDYGDTWLRHLGAPQYALGAPGATFSGYLYVADADADAPARG-PAAPPEPRSAFSA 59
QY 167 RTRSLVGSADGPGPRHQLAFDTEFVNIGDPAAGVFRCLPGAYFFSFTLGLPRKT 226
DB 60 RTRSLVGSADGPGPRHQLAFDTEFVNIGDPAAGVFRCLPGAYFFSFTLGLPRKT 119
QY 227 LSVKLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRGDAVWLLSHDHGCGAYSNHGK 286
DB 120 LSVKLMKNRDEVQAMTYDDGASRRREMOSQSVRLPLRGDAVWLLSHDHGCGAYSNHGK 179
QY 287 YITFSGFLVPPDLAPAPGLGASEL 312
DB 180 YITFSGFLVPPDLAPAPGLKPEL 205

RESULT 5
Q9H667 PRELIMINARY; PRT; 158 AA.
AC Q9H667;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ22569.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isogai T., Sugano S.;
RT "NEO human cDNA sequencing project.";
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
[2]

SEQUENCE FROM N.A.
RC TISSUE=uterus;
RA Strausberg R.;
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK026222; BAB15398.1; -;
DR EMBL; BC007520; AA07520.1; -;
DR InterPro; IPR001073; Clq.
DR Pfam; PF00386; Clq; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Hypothetical protein.
SQ SEQUENCE 158 AA; 17625 MW; 47DB10EDD6DC9760 CRC64;

Query Match 14.0%; Score 230; DB 4; Length 158;
Best Local Similarity 38.4%; Pred. No. 4.7e-12;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;
QY 147 DAPARGP-----PAPPEPRSAFSAARTSLVGSADGPGPRHQLAFDTEFVNIG 196
DB 5 DVPVTNPATILPVHVYPLPQGMVAFSAART-----SNLAPGTLDQPIVFDDLNLNIG 59
QY 197 DFDAAGVFRCLPGAYFFSFTLGLP-RKTLVKLMKNRDEVQAMTYDDGASRRREMOS 255
DB 60 TFDLQLGRFNCVPNGTYVFIFHMLKLVAVVPLVNLKNEEVLVSAYANDGAP-DHETAS 118
QY 256 QSVMLALRGDAVWLLSHDHGCGA-YSNHGKYITFSGFLVPPD 298
DB 119 NHAITQLFGDQIWLRLH---RGAITYGSSWKYSTFSGYLLYQD 158

RESULT 6
Q8TE71 PRELIMINARY; PRT; 1077 AA.
AC Q8TE71;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE EEG1L.
GN EEG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Aerdajinal W., Miller J.L.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY074490; AAL71549.1; -;
DR InterPro; IPR001073; Clq.
DR Pfam; PF00386; Clq; 1.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
SQ SEQUENCE 1077 AA; 120974 MW; 2B88BF3C47D032D6 CRC64;

Query Match 14.0%; Score 230; DB 4; Length 1077;
Best Local Similarity 38.4%; Pred. No. 5.9e-11;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;
QY 147 DAPARGP-----PAPPEPRSAFSAARTSLVGSADGPGPRHQLAFDTEFVNIG 196
DB 924 DVPVTNPATILPVHVYPLPQGMVAFSAART-----SNLAPGTLDQPIVFDDLNLNIG 978
QY 197 DFDAAGVFRCLPGAYFFSFTLGLP-RKTLVKLMKNRDEVQAMTYDDGASRRREMOS 255
DB 979 TFDLQLGRFNCVPNGTYVFIFHMLKLVAVVPLVNLKNEEVLVSAYANDGAP-DHETAS 1037
QY 256 QSVMLALRGDAVWLLSHDHGCGA-YSNHGKYITFSGFLVPPD 298
DB 1038 NHAITQLFGDQIWLRLH---RGAITYGSSWKYSTFSGYLLYQD 1077

RESULT 7

```

Q8K110
ID Q8K110 PRELIMINARY; PRT; 158 AA.
AC Q8K110;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)
DE Similar to hypothetical protein FLJ22569.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RF TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027523; AAH27523.1; -.
DR InterPro; IPR001073; Clq.
DR Pfam; PF00386; Clq; 1.
DR PRINTS; PRO0007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Hypothetical protein.
SQ SEQUENCE 158 AA; 17533 MW; 86E9321C99225FCB CRC64;

```

| | | | | |
|-----------------------|------------------|------------------|------------|-------------|
| Query Match | 13.8% | Score 226; | DB 11; | Length 158; |
| Best Local Similarity | 39.1%; | Pred. No. 1e-11; | | |
| Matches 63; | Conservative 18; | Mismatches 68; | Indels 12; | Gaps 5; |

```
QY      140 VYADADADAPARGPAPRPEPRSAFSAARTSLVGS DAGPGRHOPLAFDTEFVNIGGDED 1999
          | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      8 VTSPAAAILPVHIYPPLPOOMRVAFSART-----SNLAPGTLDQPIVFDDLNNIGETFN 62
```

QY 200 AAAGVFCRLPGAYFFSFTLGLP-RKTLISVKLMKNRDEVOAMIYDDGASRRREMOQSIV 258
| : | | | | : | : | : | : |
Db 63 LQGRFNCVPNGTYVEIFHMLKLAVNPLLYNLIMKNEEVLVSAYANDGAP-DHEIASNHAA 121

QY 259 MLALRRGDAVWLSHDHGYGA-VSNHGKYITFSGFLVPD 298
:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 122 VLQLLGDDQIWLRLH---RGAIYSSSWKYSTFSGYLLPD 158

RESULT 8

ID Q8R1P2 PRELIMINARY; PRT; 182 AA.

DT 01-JUN-2002 (TREMBlrel. 21, Created)

| | |
|----|---|
| DT | 01-JUN-2002 (Tremblrel. 21, Last sequence update) |
| DT | 01-MAR-2003 (Tremblrel. 23, Last annotation update) |

DE Similar to RIKEN cDNA 1600017K21 gene (Fragment) .
GN 1600017K21RIK.

05 Mus musculus (Mouse).
0C Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
0C Mammalia; Dumbbells; Craniata; Vertebrata; Euteleostomi;
0C Mammalia; Dumbbells; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
PN (1)

RN (1)
RP SEQUENCE FROM N.A.
RC TTSUE-Liver.

Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR MGD: 1919254; 1600017K21R1X.
DR EMBL; BC023468; AAH23468.1; -.
DR SUMMITRED (FEB-2002) CO CNE EMBL

DR InterPro; IPR001073; Clq.
DR InterPro; IPR000087; Collagen.

```
DR Pfam; PF00386; Clq; 1.
DR SMART; SM00110; Clq; 1.
```

| | | | | |
|----|----------|---------|-----------|-------------------------|
| FT | NON TER | 1 | 1 | |
| SQ | SEQUENCE | 182 AA; | 20863 MW; | 559C73DE9517882F CRC64; |

Query Match 13.5%; Score 220.5; DB 11; Length 182;

Best Local Similarity 39.2%; Pred. No. 3.7e-11;
Matches 56; Conservative 20; Mismatches 58; Indels 9; Gaps 4;

[illegible]

| | |
|-----------|--------------------------|
| RESULT 9 | . |
| Q9QXP7 | |
| ID Q9QXP7 | PRELIMINARY; PRT; 281 AA |

DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Putative secreted protein ZS1G37 (1600017K21RIK protein);
GN ZS1G37 OR 1600017K21RIK.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
 OX NCBI TaxID=10090;

RP SEQUENCE FROM N.A.
RA Sheppard P., Deisher T., Grant F., Chen L., Haldeman B., McKnight G.,
RA Whitmore T., O'Hara P.;
RT "Mus musculus putative secreted protein."
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Placenta;
RC MEDLINE=21085660; PubMed=11217851;
RX

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,

RA Kuehl P., Lewis S.

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyo-oka K., Wang K.H., Weltz C., Whittaker C., Wilmshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsuki S.,

RT "Functional annotation of a full-length mouse cDNA collection." ;
RA Hayashizaki Y. ;
PI Nature 408:685-690(2001)

RL Natute 409:685-690(2001).
DR EMBL; AF192499; AAF06664.
DR EMBL: AK005484; BA824070.

DR EMBL; AA003484; BAB24070.1; -
DR MGD; MG1:1919254; 1600017K21Rik.
DR InterPro: IPR001073; C1g.

DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.

DR PFAM; PF01391; COLLAGEN; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.

DR SMART; SM00110; C1Q; 1.
SQ SEQUENCE 281 AA; 32009 MW;

Query Match 13.5%; Score 220.5; DB 11; Length 281;

Best Local Similarity 39.2%; Pred. No. 6.5e-11;
Matches 56; Conservative 20; Mismatches 58; Indels 9; Gaps 4;

161 S A F S A R T R S L V G S D A G P G B P H O P L A F D T E F V N I G D F D A A G V F R C R L P G A Y F F S F T L G 220

Db 146 AAFSVGRKKALHSND-----YFQPVVFDTFEFVNIYKHFNMTGKFCYCVYPGIYFFSLNVH 200
QY 221 KLPRKTLVSKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRGGDAVW--LLSHDHGX 278
Db 201 TWNOKETYLHIMKNEEV-VILYAQ-VSDRSIMOSQSLMELREDEVWVRLFRGERENA 258
QY 279 GAYSNHGKYITFSGFLVYPDLAP 301
Db 259 IFSDEPDTYITFSGYLVPASEP 281

RESULT 10

ID Q8N6P2 PRELIMINARY; PRT; 243 AA.
AC Q8N6P2;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE C1q and tumor necrosis factor related protein 5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC029485; AAH29485.1; -
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
SQ SEQUENCE 243 AA; 25326 MW; 6D9306A0EB21B44A CRC64;

Query Match 13.0%; Score 213; DB 4; Length 243;
Best Local Similarity 34.9%; Pred. No. 2.4e-10;
Matches 66; Conservative 18; Mismatches 83; Indels 22; Gaps 6;

QY 122 GAPHYALGAPGATFSG---YLVDADAD---ADAPARGPAPP-----EPRSAFSAARTR 169
Db 57 GAP----GAPGKEGEGRPGLPGPRGDPGRGEGAGPAGTGPAGECSVPPRSAPSAKRSE 112
QY 170 SLVGSADGPGPRHQPLAFDTEFVNIGDFDAAGVFRCLPGAYFFSFTLGKLPKTLVS 229
Db 113 SRV-----PPPSDAPLPFDRLVNEQGHYDAVTGKTCQVPGVYFPA-VHATVYRASLQF 166
QY 230 KLKKNRDEVQAMITYDDGASRRREMOSQSVMLALRGGDAVWLLSHDHGXGAYSNHGKYIT 289
Db 167 DLVKNGESIASFFQFGGWPKPASLSGAMVRLPEPDQVWVGVDYIGIYASIKTDST 226
QY 290 FSGFLVYPD 298
Db 227 FSGFLVYSD 235

RESULT 11

ID Q8R002 PRELIMINARY; PRT; 243 AA.
AC Q8R002;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Similar to DKFZP586B0621 protein (Hypothetical 25.4 kDa protein).
GN C1QTNF5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC023068; AAH23068.1; -
DR EMBL; BC025174; AAH25174.1; -
DR MGD; MGI:2385958; C1qtnf5.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR SMART; SM00110; C1Q; 1.
KW Hypothetical protein.
SQ SEQUENCE 243 AA; 25420 MW; 498129CD051DB97B CRC64;

Query Match 12.5%; Score 205; DB 11; Length 243;
Best Local Similarity 32.8%; Pred. No. 1.2e-09;
Matches 64; Conservative 16; Mismatches 81; Indels 34; Gaps 6;

QY 122 GAPHYALGAPGATFSGYLVDADADADA-RGPPAP-----PEPRSAF 163
Db 57 GAP----GAPGKEGEG-----GRPGLPGPRGEPGRGEGAPMGAGPAGECSPPRSAF 106
QY 164 SAARTSLVGSADGPGPRHQPLAFDTEFVNIGDFDAAGVFRCLPGAYFFSFTLGKLP 223
Db 107 SAKRSESERV-----PPADTLPFDRLVNEQGHFDPTTGKTCQVPGVYFPA-VHATVY 160
QY 224 RKTLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRGGDAVWLLSHDHGXGAYSN 283
Db 161 RASLQFDLVKNGOSIASFFQYFGGWPKPASLSGAMVRLPEPDQVWVGVDYIGIYAS 220
QY 284 HGKYITFSGFLVYPD 298
Db 221 IKTDSTFSGFLVYSD 235

RESULT 12

ID Q8K479 PRELIMINARY; PRT; 243 AA.
AC Q8K479;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Complement-c1q tumor necrosis factor-related protein.
GN C1QTNF5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22135657; Pubmed=12140190;
RA Kameya S., Hawes N.L., Chang B., Heckenlively J.R., Naggert J.K., Nishina P.M.;
RT "Mfrp, a gene encoding a frizzled related protein, is mutated in the mouse retinal degeneration 6.";
RL Hum. Mol. Genet. 11:1879-1886(2002).
DR EMBL; AF469650; AAM89217.1; -
DR MGD; MGI:2385958; C1qtnf5.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 2.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
SQ SEQUENCE 243 AA; 25436 MW; 9F4D5804349791D9 CRC64;

Query Match 12.3%; Score 202; DB 11; Length 243;
Best Local Similarity 32.3%; Pred. No. 2.1e-09;
Matches 63; Conservative 17; Mismatches 81; Indels 34; Gaps 6;


```

Oy 122 GAPHYALGAPGATFSGGLVYADADADADA-RCPPAP-----PEERSAF 163
Db 57 GAP---GAPGEKEGEG-----GRPGLPGRGEPGRGEAGPMGAIGPAGECSVPPRSAP 106
Oy 164 SAARTSLVSGSDAGPGRRHQPLAFDTEFPVNIGGDFEPAAGVFRCLPLGAYFFSFTLGKLP 223
Db 107 SAKRSESERV-----PPADTPLPFDRVLLNEQGHHYDPTTGKFTCQVPGVYYFA-VHATVY 160
Oy 224 RKTLSVKLMKNRDEVOAMIVDDGASRRREMOSQSVMLALRRGDVWMLSHDHGCGAYSN 283
Db 161 RASLQFDLVKNGQSIAFFQYFGGWPKPASLSCGAWVRLPEPDQVWVGVDYIGIYAS 220
Oy 284 HGKYITFSGFLVYPD 298
Db 221 IKTDSTFSGFLVYSD 235

RESULT 13
OQ9DBU4 ID OQ9DBU4 PRELIMINARY; PRT; 294 AA.
AC OQ9DBU4;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE 1810033K05RIK protein (RIKEN CDNA 1810033K05 gene).
GN 1810033K05RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schraml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guelincich S., Hill D., Hofmann M., Hume D.A., Kamilya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayaashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Strauberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK007683; BAB25187.1; -.
DR EMBL; BC030324; AAH30324.1; -.
DR MGD; MGI:1916433; 1810033K05RIK.
DR InterPro; IPR001073; Clq.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; Clq; 1.
DR Pfam; PF01391; Collagen; 2.
DR PRINTS; PRO0007; COMPLEMENTC1Q.
DR PROSITE; PS01113; Clq; 1.
SQ SEQUENCE 294 AA; 30865 MW; 6D3905AE7C19E6FA CRC64;

Query Match 12.2%; Score 200; DB 11; Length 294;
Best Local Similarity 34.9%; Pred. No. 4e-09;
Matches 53; Conservative 30; Mismatches 45; Indels 24; Gaps 8;

```

```
QY      5 GPG---GSSELRSAFSA--RTTPLEGTSEMAVTFDKVYNIGGDFDVATGQFRCRVPGA 59
          ||| :||| :||| :||| :||| :||| :||| :|||
DB     147 PGPSCSGSSRAKSAFVAVTKSYPRE---RLPIKFDKILMNEGGHYNASSGCKFCVPGI 203
QY      60 YFESETAGAPHKSLSVMLVRNRDEVOALAFDEQRPGARRAASQSAMLQLDYGDTWLR 119
          |:|: |::| ::||| :|: |::| ::||| :|: |::| ::||| :|: |::| ::|||
DB     204 YYFTYDITLA-NKHILAIGLVHN-GQYRIRTFD-ANTGNHDVASGSTLLALKEGEDEVWLQ 259
QY      120 LHGA-----PHYALGAPGATFSGYLVAD 143
          : : : : : : : : : : : : : : : : : :
DB     260 IFYSQNGLFYDPYWT---DSLFTGFLLIYAD 287

RESULT 14
Q95JUD7 PRELIMINARY; PRT; 243 AA.
ID Q95JD7 AC Q95JUD7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Adiponectin.
SN APM1.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adipose tissue;
RX MEDLINE=21232234; PubMed=11334417;
RA Hotta K., Funahashi T., Bodkin N.L., Ortmeier H.K., Arlt Y.,
RA Hansen B.C., Matsuzawa Y.;
RT "Circulating concentrations of the adipocyte protein adiponectin are
RT decreased in parallel with reduced insulin sensitivity during the
RT progression to type 2 diabetes in rhesus monkeys.";
RL Diabetes 50:1126-1133(2001).
RD EMBL; AF404407; AAK92202.1; -.
RD InterPro; IPR001073; Clq.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; Clq; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTClQ.
DR ProDom; PD000007; Collagen; 1.
DR SMART; SM00110; ClQ; 1.
DR PROSITE; PS01113; ClQ; 1.
KM Collagen.
SQ SEQUENCE 243 AA; 26264 MW; 49A45DAF2B4613FD CRC64;

Query Match 11.8%; Score 193; DB 6; Length 243;
Best Local Similarity 31.3%; Pred. No. 1.2e-08;
Matches 47; Conservative 30; Mismatches 55; Indels 18; Gaps 6

QY      5 PGPGSELRSAFSARTTPLEGTSEMAVTFDKVYNIGGDFDVATGQFRCRVGAYFFSF 64
          ||| :||| :||| :||| :||| :||| :||| :|||
DB     103 PEGAYVYRSAFVSGLETYTV-VENMPDIRFKIFYNQNHSDSTGRHCNIPLYYFAY 161
QY      65 TAGKAPH-----KSLSVMLVRNRDEVOALA.FDEQRPGARRAASQSAMLQLDYGD TWLR 119
          | :|: |::| ::||| :|: |::| ::||| :|: |::| ::||| :|: |::| ::|||
DB     162 -----HITVMKDVKVSLFK-KDKAMLFITYDQYQENNVDQ-AAGSYVLHLHEVGQVWLQ 213

QY      120 LHGAPH----YALGAPGATFSGYLVADAD 145
          ::|: || :||| :||| :||| :||| :||| :|||
DB     214 VYGEGERNGLYADNDNDSTFTGFLIYHDTN 243

RESULT 15
Q8BRW2 PRELIMINARY; PRT; 247 AA.
ID Q8BRW2 AC Q8BRW2;
AC Q8BRW2;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
```

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Adipocyte complement related protein of 30 kDa.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL, AK041214; BAC30866.1; ..
SQ SEQUENCE 247 AA; 26751 MW; 0D3FA64C789CAEF3 CRC64;

| | | | | | |
|-----------------------|--|--------------|--------------------|------------|-----------------------|
| Query Match | | 11.5%; | Score 189; | DB 11; | Length 247; |
| Best Local Similarity | | 33.5%; | Pred. No. 2.8e-08; | | |
| Matches | 53; | Conservative | 24; | Mismatches | 73; Indels 8; Gaps 5; |
| QY | 146 ADAPARGPPAPP---EPRSAFSAARTRSLVGSDDAGPGRHOPLAFDTEFVNIGDFDAA 201 | | | | |
| DB | 91 AEGP-RGFPGTGRKGEPEGEAAVYRSAFSVGLETRVTVPNVPIRFTKIFYNQNHYDGS 149 | | | | |
| QY | 202 AGVFCRLPGAYFFSFTLGKLPRKTLVKLMKNRDEVQAMTYDDGASRRREMOSQSVMLA 261 | | | | |
| DB | 150 TGKFYCNIPGLYFYSYHI-TVYMKDVKVSLEFR-KDKAVLFTYDQYQEKNVDDQASGSVLLH 207 | | | | |
| QY | 262 LRRGDAVWLISH-DHDGYGAYSNHGKYYTTFSGFLVYPD 298 | | | | |
| DB | 208 LEVGDQVWLQVYGDGDHNGLYADNVNDSTFTGFLLYHD 245 | | | | |

Search completed: January 12, 2004, 08:19:08
Job time : 44.1078 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:09 ; Search time 47.1687 Seconds

(without alignments)
1053.272 Million cell updates/sec

Title: US-10-085-167-2_COPY_17_329

Perfect score: 1639

Sequence: 1 LGPTPGSGSELRSFAFSAAR.....LVYPDLAPAAPGLGASELL 313

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

1: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|----------|--------------------|
| 1 | 1639 | 100.0 | 329 | 22 | AAB61606 | Human ZACRP4. Hom |
| 2 | 1620 | 98.8 | 329 | 23 | ABG79643 | Human novel secret |
| 3 | 1579 | 96.3 | 329 | 22 | AAB61424 | Monkey MANGO 245 p |
| 4 | 1480 | 90.3 | 348 | 22 | AAB61423 | Human MANGO 245 pr |
| 5 | 1476 | 90.1 | 334 | 22 | AAB61466 | Human MANGO 245 ma |
| 6 | 1453 | 88.7 | 299 | 23 | ABG70385 | Adipocyte compleme |
| 7 | 1358.5 | 82.9 | 284 | 23 | ABG70384 | Adipocyte compleme |
| 8 | 809 | 49.4 | 199 | 22 | AAB61488 | Murine MANGO 245 p |
| 9 | 760 | 46.4 | 192 | 22 | AAB61479 | Human MANGO 245 cy |

| | | | | | | |
|----|-------|------|------|----|----------|--------------------|
| 10 | 707.5 | 43.2 | 221 | 23 | ABG70383 | Adipocyte compleme |
| 11 | 683 | 41.7 | 134 | 22 | AAB61473 | Monkey MANGO 245 C |
| 12 | 664 | 40.5 | 127 | 22 | ABB11578 | Human secreted pro |
| 13 | 646 | 39.4 | 126 | 22 | AAB61469 | Human MANGO 245 C1 |
| 14 | 639 | 39.0 | 126 | 22 | AAB61489 | Murine MANGO 245 C |
| 15 | 635 | 38.7 | 125 | 22 | AAB61477 | Human MANGO 245 ex |
| 16 | 622 | 37.9 | 126 | 22 | AAB61472 | Monkey MANGO 245 C |
| 17 | 618 | 37.7 | 123 | 21 | AAB42189 | Human ORFX ORF1953 |
| 18 | 609 | 37.2 | 117 | 22 | AAB61470 | Human MANGO 245 C1 |
| 19 | 552 | 33.7 | 130 | 22 | AAB61468 | Mature monkey MANG |
| 20 | 507 | 30.9 | 133 | 23 | ABG70386 | Adipocyte compleme |
| 21 | 237.5 | 14.5 | 252 | 22 | AAB49590 | Human zacrp5. Hom |
| 22 | 237.5 | 14.5 | 252 | 23 | AAO19412 | Human molecule for |
| 23 | 237.5 | 14.5 | 252 | 23 | AAU83100 | Novel secreted pro |
| 24 | 237.5 | 14.5 | 252 | 23 | ABB53292 | Human polypeptide |
| 25 | 231 | 14.1 | 800 | 22 | AAU78334 | Human protein SEQ |
| 26 | 230 | 14.0 | 202 | 22 | AAU99927 | Human polypeptide |
| 27 | 230 | 14.0 | 709 | 23 | ABB08645 | Human pancreatic C |
| 28 | 230 | 14.0 | 710 | 23 | AAO15420 | Human gensef metab |
| 29 | 230 | 14.0 | 746 | 20 | AAU29512 | Human lung tumour |
| 30 | 230 | 14.0 | 746 | 21 | AAB44461 | Human lung tumour |
| 31 | 230 | 14.0 | 746 | 22 | AAE13802 | Human lung tumour |
| 32 | 230 | 14.0 | 908 | 22 | AAE09839 | Novel human protei |
| 33 | 230 | 14.0 | 909 | 22 | AAE09845 | Novel human protei |
| 34 | 230 | 14.0 | 957 | 22 | AAE09841 | Novel human protei |
| 35 | 230 | 14.0 | 958 | 22 | AAE09847 | Novel human protei |
| 36 | 230 | 14.0 | 992 | 22 | AAE09843 | Novel human protei |
| 37 | 230 | 14.0 | 993 | 22 | AAE09849 | Novel human protei |
| 38 | 230 | 14.0 | 1042 | 22 | AAE09838 | Novel human protei |
| 39 | 230 | 14.0 | 1043 | 22 | AAE09844 | Novel human protei |
| 40 | 230 | 14.0 | 1091 | 22 | AAE09840 | Novel human protei |
| 41 | 230 | 14.0 | 1092 | 22 | AAE09846 | Novel human protei |
| 42 | 230 | 14.0 | 1126 | 22 | AAE09842 | Novel human protei |
| 43 | 230 | 14.0 | 1127 | 22 | AAE09848 | Novel human protei |
| 44 | 220.5 | 13.5 | 281 | 20 | AAU98014 | Mouse adipocyte-sp |
| 45 | 220.5 | 13.5 | 281 | 21 | AAB08428 | Amino acid sequenc |

ALIGNMENTS

| | |
|----------|--|
| RESULT 1 | |
| AAB61606 | |
| ID | AAB61606 standard; Protein; 329 AA. |
| XX | |
| AC | AAB61606; |
| XX | |
| DT | 05-APR-2001 (first entry) |
| XX | |
| DE | Human ZACRP4. |
| XX | |
| KW | Human; zacrp4; complement factor C1q domain; chromosome 11q11; |
| KW | energy balance; cellular metabolic reaction; autocrine factor; |
| KW | development; cell proliferation; differentiation; cell survival. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| EH | |
| FT | Key |
| FT | Peptide |
| FT | Domain |
| FT | Domain |
| FT | Domain |
| FT | Domain |
| XX | |
| PN | WO200102565-A2. |
| XX | |
| PD | 11-JAN-2001. |
| XX | |
| PF | 28-JUN-2000; 2000WO-US17692. |
| XX | |
| PR | 01-JUL-1999; 99US-0346502. |
| XX | |

Location/Qualifiers
1..16
/label= Signal_peptide
17..159
/label= C1q_domain_#1
160..328
/label= C1q_domain_#2

PA (ZYMO) ZYMOGENETICS INC.
XX
PI Holloway JL, Lok S;
XX
DR WPI; 2001-138140/14.
DR N-PSDB; AAF28672.
XX
PT Novel secreted protein ZACRP4 polypeptides having tandem C1q globular
PT domains, useful for studying cell-cell communication and regulation of
PT cellular processes -
XX
PS Claim 1; Page 77-78; 82pp; English.
XX
CC The present sequence is human ZACRP4 protein. ZACRP4 protein has two
CC complement factor C1q domains. The ZACRP4 gene is located on human
CC chromosome 11q11. The ZACRP4 coding sequence and protein have a number of
CC uses described in the specification, including, modulation of energy
CC balance and cellular metabolic reactions in mammals. In addition, ZACRP4
CC protein is useful as an autocrine factor, particularly during
CC development, in mediating the processes of an organism, in regulating
CC cellular processes such as cell proliferation and/or differentiation,
CC cell survival and energy balance.
XX
SQ Sequence 329 AA;

Query Match 100.0%; Score 1639; DB 22; Length 329;
Best Local Similarity 100.0%; Pred. No. 4e-152;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LGPTPGSGSELRSASFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAY 60
DB 17 LGPTPGSGSELRSASFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAY 76
OY 61 FFSFTAGKAPHKSLSVMLVRNRDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTWRL 120
DB 77 FFSFTAGKAPHKSLSVMLVRNRDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTWRL 136
OY 121 HGAPHYALGAPGATFSGYLVYADADADAPARGPPAPPEPRSAFSAARTSLVGS DAGPGP 180
DB 137 HGAPHYALGAPGATFSGYLVYADADADAPARGPPAPPEPRSAFSAARTSLVGS DAGPGP 196
OY 181 RHQPLAFDTEFVNIIGDFDAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQA 240
DB 197 RHQPLAFDTEFVNIIGDFDAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQA 256
OY 241 MIYDDGASRRREMOSQSVMLALRGGDAVWLISHDHGCGAYSNHGKXITTFSGFLVYPPDLA 300
DB 257 MIYDDGASRRREMOSQSVMLALRGGDAVWLISHDHGCGAYSNHGKXITTFSGFLVYPPDLA 316
OY 301 PAAPPGLGASELL 313
DB 317 PAAPPGLGASELL 329

RESULT 2
ABG79643
ID ABG79643 standard; Protein; 329 AA.

XX
AC ABG79643;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human novel secreted protein SECP19, Incyte ID No. 931619CD1.
XX
KW Human; SECP; secreted protein; cell proliferative disorder;
KW actinic keratosis; arteriosclerosis; bursitis; hepatitis; cancer;
KW autoimmune disorder; inflammatory disorder; AIDS; asthma; allergy;
KW acquired immunodeficiency syndrome; anaemia; atopic dermatitis;
KW cardiovascular disorder; congestive heart failure; vascular tumour;
KW ischaemic heart disease; myocardial infarction; epilepsy; stroke;
KW hypertensive heart disease; neurological disorder; cerebral neoplasm;
KW Alzheimer's disease; developmental disorder; renal tubular acidosis;
KW Cushing's syndrome; Duchenne muscular dystrophy; hypothyroidism;

KM Becker muscular dystrophy.
XX
OS Homo sapiens.
XX
PN WO200262841-A2.
XX
PD 15-AUG-2002.
XX
PF 28-JAN-2002; 2002WO-US02616.
XX
PR 02-FEB-2001; 2001US-266195P.
PR 08-FEB-2001; 2001US-267924P.
PR 09-FEB-2001; 2001US-267816P.
PR 09-FEB-2001; 2001US-268112P.
PR 26-FEB-2001; 2001US-271639P.
PR 07-SEP-2001; 2001US-317818P.
PR 21-DEC-2001; 2001US-343553P.
XX
PA (INCYTE GENOMICS INC.
XX
PI Tang TY, Yue H, Gandhi AR, Yao MG, Warren BA, Ding L, Duggan BM;
PI Xu Y, Yang J, Thangavelu K, Lai PG, Honchell CD, Walla NK, Lee S;
PI Lee EA, Richardson TW, Baughn MR, Elliott VS;
XX
DR WPI; 2002-657522/70.
DR N-PSDB; ABS64954.

PT New human secreted proteins and nucleic acids useful in diagnosing,
PT treating and preventing cell proliferative, autoimmune/inflammatory,
PT cardiovascular, neurological, and developmental disorders -
XX
PS Claim 1; Page 140; 158pp; English.

CC The invention relates to twenty four human secreted proteins
CC (SECP1-24), proteins 90% identical to them and active fragments of them.
CC Also included are nucleic acids encoding the SECP proteins, a recombinant
CC polynucleotide comprising a promoter sequence operably linked to the
CC nucleic acid, a cell transformed with the recombinant polynucleotide, an
CC transgenic organism comprising the recombinant polynucleotide, an
CC anti-SECP antibody, and screening for ant/agonists and modulators of
CC SECP function or expression. The SECP proteins and nucleic acids are
CC useful in the diagnosis, treatment and prevention of cell proliferative
CC (e.g. actinic keratosis, arteriosclerosis, bursitis, hepatitis or
CC cancer), autoimmune/inflammatory (e.g. AIDS (acquired immunodeficiency
CC syndrome), asthma, anaemia, allergies or atopic dermatitis),
CC cardiovascular (e.g. congestive heart failure, ischaemic heart disease,
CC myocardial infarction, hypertensive heart disease, or vascular tumours),
CC neurological (e.g. epilepsy, stroke, cerebral neoplasms, or Alzheimer's
CC disease), and developmental (e.g. renal tubular acidosis, Cushing's
CC syndrome, Duchenne and Becker muscular dystrophy, or hypothyroidism)
CC disorders. Many other diseases and disorders are listed in the
CC specification. These may also be used in assessing the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid
CC SECP protein of the invention.

SQ Sequence 329 AA;

Query Match 98.8%; Score 1620; DB 23; Length 329;
Best Local Similarity 99.4%; Pred. No. 2.9e-150;
Matches 311; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 LGPTPGSGSELRSASFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAY 60
DB 17 LGPTPGSGSELRSASFSAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAY 76
OY 61 FFSFTAGKAPHKSLSVMLVRNRDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTWRL 120
DB 77 FFSFTAGKAPHKSLSVMLVRNRDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTWRL 136
OY 121 HGAPHYALGAPGATFSGYLVYADADADAPARGPPAPPEPRSAFSAARTSLVGS DAGPGP 180
DB 137 LGAPHYALGAPGATFSGYLVYADADADAPARGPPAPPEPRSAFSAARTSLVGS DAGPGP 196

QY 181 RHQPLAFDTEFVNIGDFDAAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQA 240
Db 197 RHQPLAFDTEFVNIGDFDAAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQA 256
QY 241 MIYDDGASRRREMOSQSVMLALRRGDAVWLSSHHDGYGAYSNHGKYITFSGLVYPDLA 300
Db 257 MIYDDGASRRREMOSQSVMLALRRGDAVWLSSHHDGYGAYSNHGKYITFSGLVYPDLA 316
QY 301 PAAPPGLGASELL 313
Db 317 PAAPPGLGASELL 329

RESULT 3
AAB61424
ID AAB61424 standard; protein; 329 AA.
XX
AC AAB61424;
XX
DT 04-APR-2001 (first entry)
XX
DE Monkey MANGO 245 protein.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Catarrhini sp..
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX

PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Claim 1; Fig 25; 262pp; English.
XX

CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 329 AA;

Query Match 96.3%; Score 1579; DB 22; Length 329;
Best Local Similarity 96.5%; Pred. No. 3.1e-146;
Matches 302; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 LGPTPGGSSSELRSASFSAARTTPLEGTSEMAVTFDKVYVNIGDFDVATGQFRCRVPGAY 60
Db 17 LGPAPGPGSSSELRSASFSAARTTPLEGASEMAVTFDKVYVNIGDFDAATGQFRCRVPGAY 76
QY 61 FFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMQLDYGDYTWLRL 120
Db 77 FFSFTVGAAPHKSLSVMLVRNHDVEQALAFDEQRRPSARRAASQSAMQLDYGDYTWLRL 136

QY 121 HGAPHYALGAPGATFSGYLVYADADADAPARGPPAPPEPRSAFSAARTSLVGS DAGPGP 180
Db 137 HGAPQYALGAPGATFSGYLVYADADADAPARGPPAPPEPRSAFSAARTSLVGS DAGSGP 196
QY 181 RHQPLAFDTEFVNIGDFDAAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQA 240
Db 197 RHQPLAFDTELVNIGDFDAAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQA 256
QY 241 MIYDDGASRRREMOSQSVMLALRRGDAVWLSSHHDGYGAYSNHGKYITFSGLVYPDLA 300
Db 257 MIYDDGASRRREMOSQSVMLALRRGDAVWLSSHHDGYGAYSNHGKYITFSGLVYPDLA 316
QY 301 PAAPPGLGASELL 313
Db 317 PAAPPGLGAPPELL 329

RESULT 4
AAB61423
ID AAB61423 standard; protein; 348 AA.
XX
AC AAB61423;
XX
DT 04-APR-2001 (first entry)
XX
DE Human MANGO 245 protein.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Homo sapiens.
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX

PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Claim 1; Fig 23; 262pp; English.
XX

CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 348 AA;

Query Match 90.3%; Score 1480; DB 22; Length 348;
Best Local Similarity 99.6%; Pred. No. 1.7e-136;
Matches 283; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LGPTPGGSSSELRSASFSAARTTPLEGTSEMAVTFDKVYVNIGDFDVATGQFRCRVPGAY 60
Db 17 LGPTPGGSSSELRSASFSAARTTPLEGTSEMAVTFDKVYVNIGDFDVATGQFRCRVPGAY 76

| | | | |
|----|-----|--|-----|
| Qy | 61 | FFSFTAGKAPHKSLSVMLVRNRDEVQALAFDEQRRPGARRAASQSAMLQDLDPYDVTWLR | 120 |
| Db | 77 | FFSFTAGKAPHKSLSVMLVRNRDEVQALAFDEQRRPGARRAASQSAMLQDLDPYDVTWLR | 136 |
| Qy | 121 | HGAPHYALGAPGATFSGYLVYADADADADAPARGPPAPPEPRSAFSAARTRSLVGS DAGPP | 180 |
| Db | 137 | HGAPQYALGAPGATFSGYLVYADADADADAPARGPPAPPEPRSAFSAARTRSLVGS DAGPP | 196 |
| Qy | 181 | RHQPLAFDTEFVNIGGDFDAAGVFRCRLPBAYFFSFTLGLPRKTL SVKLMKNRDEVQ | 240 |
| Db | 197 | RHQPLAFDTEFVNIGGDFDAAGVFRCRLPBAYFFSFTLGLPRKTL SVKLMKNRDEVQ | 256 |
| Qy | 241 | MIYDDGASRRREMOSQSVMLALRRGDVWLISHDHDGYGAYSNH | 284 |
| Db | 257 | MIYDDGASRRREMOSQSVMLALRRGDVWLISHDHDGYGAYSNH | 300 |

RESULT 5

| | | |
|----|----|---|
| XX | AC | AAB61466; |
| XX | DT | 04-APR-2001 (first entry) |
| XX | DE | Human MANGO 245 mature protein. |
| XX | KW | TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung; |
| XX | KW | autoimmune; allergy; cardiovascular; brain; degenerative; placental; |
| XX | KW | pancreatic; skeletal; muscle. |
| OS | | Homo sapiens. |
| XX | PN | WO200100672-A1. |
| XX | PD | 04-JAN-2001. |
| XX | PF | 29-JUN-2000; 2000WO-US18184. |
| XX | PR | 29-JUN-1999; 99US-0342687. |
| XX | PA | (MILL-) MILLENNIUM PHARM INC. |
| XX | PI | Holtzman DA, Barnes TM, Fraser CC, Sharp JD; |
| XX | DR | WPI; 2001-050127/06. |
| XX | PT | Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and |
| XX | PT | MANGO 245 proteins, useful in the treatment of inflammatory diseases |
| XX | PT | (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver |
| XX | PT | disorders (e.g. jaundice) - |
| XX | PS | Disclosure; Page 242; 262pp; English. |
| XX | CC | The present invention relates to cDNAs encoding TANGO 244, |
| XX | CC | TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins. |
| XX | CC | The nucleic acids, proteins and protein modulators are useful for |
| XX | CC | treating colonic disorders, inflammatory diseases, tumors, |
| XX | CC | renal disorders, liver disorders, lung disorders, autoimmune diseases, |
| XX | CC | allergic diseases, cardiovascular diseases, brain disorders, |
| XX | CC | degenerative diseases placental, pancreatic, skeletal and muscle |
| XX | CC | disorders. |

| | | |
|---------------------------|---------------|--------------------------------|
| SQ | Sequence | 334 AA; |
| Query Match | 90.1%; | Score 1476; DB 22; length 334; |
| Best Local Similarity | 99.3%; | Pred. No. 3.9e-136; |
| Matches 282; Conservative | 0; Mismatches | 2; Indels 0; Gaps 0; |

| QY | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 |
|----|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Db | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 |

| | | | |
|----|-----|--|-----|
| QY | 61 | FFSFTAGKAPHKSLSVMLVNRNDEVQALAFDEQRRPGARRAASQSAMLQLDYDGTWMLRL | 120 |
| | | | |
| | | | |
| | | | |
| Db | 63 | FFSFTAGKAPHKSLSVMLVNRNDEVQALAFDEQRRPGARRAASQSAMLQLDYDGTWMLRL | 122 |
| | | | |
| | | | |
| | | | |
| QY | 121 | HGAPHYALGAPGATFSGYLVYADADADAPARGPPAPPEPRSAFSAARTRSLVGS DAGGP | 180 |
| | | | |
| | | | |
| | | | |
| Db | 123 | HGAPQYALGAPGATFSGYLVYADADADADAPARGPPAPPEPRSAFSAARTRSLVGS DAGGP | 182 |
| | | | |
| | | | |
| | | | |
| QY | 181 | RHQP LAFDTEFVNIIGSDFDAAGVFRCLPGAYFFSFTLGLPKRTL SVKLMKNRDEVQ A | 240 |
| | | | |
| | | | |
| | | | |
| Db | 183 | RHQP LAFDTEFVNIIGSDFDAAGVFRCLPGAYFFSFTLGLPKRTL SVKLMKNRDEVQ A | 242 |
| | | | |
| | | | |
| | | | |
| QY | 241 | MIYDDGASRRREMOSQSVMLALRGDAVWLLSHDHG YGAYS NH | 284 |
| | | | |
| | | | |
| | | | |
| Db | 243 | MIYDDGASRRREMOSQSVMLALRGDAVWLLSHDHG YGAYS NH | 286 |
| | | | |
| | | | |
| | | | |

RESULT 6

| | | |
|----|--|----------------------------|
| ID | ABG70385 | standard; Protein; 299 AA. |
| XX | | |
| AC | ABG70385; | |
| XX | | |
| DT | 05-NOV-2002 | (first entry) |
| XX | | |
| DE | Adipocyte complement-related C1q Tumour Necrosis Factor-like protein #3. | |
| XX | | |
| KW | Human; NOX; NOX-associated disorder; cardiomyopathy; atherosclerosis; | |
| KW | cell signal processing; metabolic pathway modulation; metabolic disorder | |
| KW | obesity; diabetes; infectious disease; neurodegenerative disorder; acne; | |
| KW | Alzheimer's disease; Parkinson's disease; immune disorder; cancer; | |
| KW | haematopoietic disorder; cirrhosis; pancreatitis; learning defect; | |
| KW | memory defect; infertility; congenital heart defect; hair growth; | |
| KW | pigmentation disorder; endocrine disorder; respiratory disease; health; | |
| KW | gastro-intestinal disease; reproductive; neurological disease; | |
| KW | bone marrow transplantation; endocrine disease; allergy; inflammation; | |
| KW | nephrological disorder; urinary system disorder; age-related disorder; | |
| KW | neuropsychiatric disorder; EGF-related protein; SCUBE1; TEN-M4; | |
| KW | adipocyte complement-related C1q tumour necrosis factor; out at first; | |
| KW | beta adrenergic receptor kinase; EphA6/enk-2; glucose transporter; | |
| KW | type Ia membrane sushi-containing domain; butyrophilin; | |
| KW | type Ia membrane-sushi domain containing. | |
| XX | | |
| OS | Homo sapiens. | |
| XX | | |
| PN | WO200257453-A2. | |
| XX | | |
| PD | 25-JUL-2002. | |
| XX | | |
| PF | 19-DEC-2001; 2001WO-US50331. | |
| XX | | |
| PR | 19-DEC-2000; 2000US-265704P. | |
| PR | 20-DEC-2000; 2000US-257314P. | |
| PR | 02-MAY-2001; 2001US-288153P. | |
| PR | 29-MAY-2001; 2001US-294075P. | |
| PR | 24-JUL-2001; 2001US-307506P. | |
| PR | 10-AUG-2001; 2001US-311590P. | |
| PR | 10-AUG-2001; 2001US-311613P. | |
| PR | 29-AUG-2001; 2001US-315617P. | |
| PR | 14-SEP-2001; 2001US-322358P. | |
| XX | | |
| PA | (CURA-) CURAGEN CORP. | |

PI Gangolli EA, Patturajan M, Vernet CAM, Malyanar UM, Kekuda R;
PI Stone DJ, Anderson D, Shimkets RA, Burgess CE, Zernusen BD, Liu X;
PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
DR WPI; 2002-590744/63.
DR N-PSDB; ABB52097.

Novel isolated NOVX polypeptide useful for treating cardiomyopathy, atherosclerosis, metabolic disorders, diabetes, obesity, infectious PT

PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
PT cancer -
XX
XX
PS Claim 1; Page 31; 318pp; English.
XX
XX The present invention relates to new NOVX polypeptides. The invention is
CC useful for treating or preventing a NOVX-associated disorder such as
CC cardiomyopathy or atherosclerosis, where the disorder is related to cell
CC signal processing and metabolic pathway modulation in a subject,
CC preferably human. The invention is also useful for treating metabolic
CC disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
CC disorders, haematopoietic disorders and various cancers. The molecules of
CC the invention are also useful for treating or preventing cirrhosis,
CC pancreatitis, learning and memory defects, infertility, congenital heart
CC defects, acne, hair growth, pigmentation disorders, endocrine disorders,
CC respiratory disease, gastro-intestinal diseases, reproductive, health,
CC neurological diseases, bone marrow transplantation, endocrine diseases,
CC allergy and inflammation, nephrological disorders, urinary system
CC disorders, neuropsychiatric disorders and age-related disorders.
CC The present amino acid sequence represents a NOVX protein of the
XX invention.
XX
XX
XX Sequence 299 AA;
SQ

| | | | | |
|---------------------------|--------|---------------------|------------|-------------|
| Query Match | 88.7%; | Score 1453; | DB 23; | Length 299; |
| Best Local Similarity | 90.1%; | Pred. No. 6.1e-134; | | |
| Matches 282; Conservative | 0; | Mismatches 1; | Indels 30; | Gaps 1; |

| | | | | | | |
|----|-----|---|----------------|-------------|-----------|------------|
| QY | 1 | LGPTPGGSSSELRSAFSAARTTPLEGTSEMAVTFDKVYVNI | GDGDFVATGQFR | CRVP | GAY | 60 |
| Db | 17 | LGPTPGGSSSELRSAFSAARTTPLEGTSEMAVTFDKVYVNI | GDGDFVATGQFR | CRVP | GAY | 76 |
| QY | 61 | FFSFTAGKAPHKSLSVMLVNRDEVQALAFDEQRRPGARRA | SQSAMLQLDY | GD | TW | 120 |
| Db | 77 | FFSFTAGKAPHKSLSVMLVNRDEVQALAFDEQRRPGARRA | SQSAMLQLDY | GD | TW | 136 |
| QY | 121 | HGAPHYALGAPGATFSGYL | VYADADADAPARGP | PAPEPRSAFSA | ARTRSLVGS | DAGPGP 180 |
| Db | 137 | HGAPHYALGAPGATFSGYL | VYADADADAPARGP | PAPEPRSAFSA | ARTRSLVGS | DAGPGP 166 |

| | | |
|----|---|-----|
| QY | 181 RHQPLAFDTEFFVNIIGDDEPAAAGVFRCLPGAYFFSFETLGLPRKTLVSKLMKNRDEVQA | 240 |
| Db | 167 RHQPLAFDTEFFVNIIGDDEPAAADVFRCLPGAYFFSFETLGLPRKTLVSKLMKNRDEVQA | 226 |
| QY | 241 MIYDDGASRRREMOSQSVMLALRRGDVAVMLLSHDHDGYGAYSNHGKIYTFSGFLVYPDLA | 300 |
| Db | 227 MIYDDGASRRREMOSQSVMLALRRGDVAVMLLSHDHDGYGAYSNHGKIYTFSGFLVYPDLA | 286 |

| | | | |
|----|-----|---------------|-----|
| QY | 301 | PAAPPGLGASELL | 313 |
| | | | |
| Db | 287 | PAAPPGLGASELL | 299 |

| | |
|----------|------------------------------------|
| RESULT 7 | |
| ABG70384 | |
| ID | ABG70384 standard; Protein; 284 AA |

AC ABG70384;

DT 05-NOV-2002 (first entry)

DE Adipocyte complement-related C1q Tumour Necrosis Factor-like protein #2.

KW Human; NOX; NOX-associated disorder; cardiomyopathy; atherosclerosis;
 KW cell signal processing; metabolic pathway modulation; metabolic disorder;
 KW obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
 KW Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
 KW haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
 KW memory defect; infertility; congenital heart defect; hair growth;
 KW pigmentation disorder; endocrine disorder; respiratory disease; health;
 KW gastro-intestinal disease; reproductive; neurological disease;
 KW bone marrow transplantation; endocrine disease; allergy; inflammation;

KW nephrological disorder; urinary system disorder; age-related disorder;
KW neuropsychiatric disorder; EGF-related protein; SCUBE1; TEN-M4;
KW adipocyte complement-related C1q tumour necrosis factor; out at first;
KW beta-adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;
KW type 1a membrane sushi-containing domain; butyrophillin;
KW type 1a membrane-sushi domain containing.

OS Homo sapiens.

PN WO200257453-A2.

PD 25-JUL-2002.

PF 19-DEC-2001; 2001WO-US50331.

PR 19-DEC-2000; 2000US-265704P.

PR 02-MAY-2001; 2001US-288153P.

PR 24-JUL-2001; 2001US-307506P.

PR 10-AUG-2001; 2001US-311613P.

PR 14-SEP-2001; 2001US-322358P.

PA (CURA-) CURAGEN CORP.

PI Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;

PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;

DR WPI; 2002-590744/63.

DR N-PSDB; ABS52096.

Novel isolated NOVX polypeptide useful for treating cardiomyopathy, PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or PT cancer -

PS Claim 1; Page 30; 318pp; English..

The present invention relates to new NOVX polypeptides. The invention is useful for treating or preventing a NOVX-associated disorder such as cardiomyopathy or atherosclerosis, where the disorder is related to cell signal processing and metabolic pathway modulation in a subject, preferably human. The invention is also useful for treating metabolic disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative disorders (e.g. Alzheimer's disease, Parkinson's disease), immune disorders, haematopoietic disorders and various cancers. The molecules of the invention are also useful for treating or preventing cirrhosis, pancreatitis, learning and memory defects, infertility, congenital heart defects, acne, hair growth, pigmentation disorders, endocrine disorders, respiratory disease, gastro-intestinal diseases, reproductive, health, neurological diseases, bone marrow transplantation, endocrine diseases, allergy and inflammation, nephrological disorders, urinary system disorders, neuropsychiatric disorders and age-related disorders. The present amino acid sequence represents a NOVX protein of the invention.

... SQ Sequence 284 AA;

| | | | | |
|-----------------------|--------------|-------------------|----------------|-------------|
| Query Match | 82.9%; | Score 1358.5; | DB 23; | Length 284; |
| Best Local Similarity | 85.3%; | Pred. No. 1e-124; | | |
| Matches 267; | Conservative | 0; | Mismatches .1; | Indels 45; |
| | | | | Gaps 1; |

[illegible]

| | | | |
|----|-----|--|-----|
| QY | 121 | HGAPHYALGAPGATFSGYLVYADADADADAPARGPPAPPEPRSAFSAARTRSLVGS DAGGP | 180 |
| | | | |
| Db | 137 | HGAPQYALGAPGATFSGYLVYADA----- | 160 |
| QY | 181 | RHQPLAFDTEFVNIGGDFDAAGVFRCRLPGAYFFSFTLGKLPRKTL SVKLMKNRDEVQA | 240 |
| | | | |
| Db | 161 | -----EFVNIGGDFDAAGVFRCRLPGAYFFSFTLGKLPRKTL SVKLMKNRDEVQA | 211 |
| QY | 241 | MITYDGASRRREMOSQSVM LALRRGD AVMLLSHDH DGYGAYSNHGKYITFSGF LVYPDLA | 300 |
| | | | |
| Db | 212 | MITYDGASRRREMOSQSVM LALRRGD AVMLLSHDH DGYGAYSNHGKYITFSGF LVYPDLA | 271 |
| QY | 301 | PAAPPGIGASELL 313 | |
| | | | |
| Db | 272 | PAAPPGIGASELL 284 | |

| | |
|----------|---|
| RESULT 8 | |
| AAB61488 | |
| ID | AAB61488 standard; protein; 199 AA. |
| XX | |
| AC | AAB61488; |
| XX | |
| DT | 04-APR-2001 (first entry) |
| XX | |
| DE | Murine MANGO 245 protein. |
| XX | |
| KW | TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung; |
| KW | autoimmune; allergy; cardiovascular; brain; degenerative; placental; |
| KW | pancreatic; skeletal; muscle. |
| XX | |
| OS | Mus musculus. |
| XX | |
| PN | WO200100672-A1. |
| XX | |
| PD | 04-JAN-2001. |
| XX | |
| PF | 29-JUN-2000; 2000WO-US18184. |
| XX | |
| PR | 29-JUN-1999; 99US-0342687. |
| XX | |
| PA | (MILL-) MILLENNIUM PHARM INC. |
| XX | |
| PI | Holtzman DA, Barnes TM, Fraser CC, Sharp JD; |
| XX | |
| DR | WPI; 2001-050127/06. |
| XX | |
| PT | Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and |
| PT | MANGO 245 proteins, useful in the treatment of inflammatory diseases |
| PT | (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver |
| PT | disorders (e.g. jaundice) - |
| XX | |
| PS | Claim 1; Fig 29; 262pp; English. |
| XX | |
| CC | The present invention relates to cDNAs encoding TANGO 244, |
| CC | TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins. |
| CC | The nucleic acids, proteins and protein modulators are useful for |
| CC | treating colonic disorders, inflammatory diseases, tumors, |
| CC | renal disorders, liver disorders, lung disorders, autoimmune diseases, |
| CC | allergic diseases, cardiovascular diseases, brain disorders, |
| CC | degenerative diseases placental, pancreatic, skeletal and muscle |
| CC | disorders. |
| XX | |
| Sequence | 199 AA; |

| | | | | |
|---------------------------|-------|--------------------|-----------|-------------|
| Query Match | 49.4% | Score 809; | DB 22; | Length 199; |
| Best Local Similarity | 87.2% | Pred. No. 4.7e-71; | | |
| Matches 163; Conservative | 4; | Mismatches 16; | Indels 4; | Gaps 3; |

```
Qy      1 LGPTPGSSSELSRASFSAARTTPLEGTSMAVTFDKYVYNIGDFFDVATGQFRCPVGAY    60
        ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     17 LGPA-GPGSSELRSASFSAARTTPLEGTSEMAVTFDKYVYNIGDFDAATGRFRRVPGAY    75
```

```

QY      61  FFSFTAGKAPHKSLSVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDYDGTWMLRL 120
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      76  FFSFTAGKAPHKNSVMLVRNDEVQALAFDQRRPGARRAASQSAMLQLDYDGTWMLRL 135
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY      121 HGAPHYALGAPGATFSGYLVYADADADADAPARGPPAPPEPRSAFSAARTSLVGS DAGGPP 180
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      136 HGAPHYALGAPGATFSGYLVYADADADADAPARG-PAAPPEPRSAFSAARHTWMAPNP RPAP 194
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY      181 RHQPLAF 187
      |||
Db      195 RR-LAF 199

```

RESULT 9
AAB61479
ID AAB61479 standard; protein; 192 AA.
XX
AC AAB61479;
XX
DT 04-APR-2001 (first entry)
XX
DE Human MANGO 245 cytoplasmic domain.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Homo sapiens.
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Disclosure; Page 247; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 192 AA;

| | | | | |
|-----------------------|-----------------|--------------------|-----------|-------------|
| Query Match | 46.4%; | Score 760; | DB 22; | Length 192; |
| Best Local Similarity | 100.0%; | Pred. No. 2.9e-66; | | |
| Matches 144; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |

[illegible]

| Qy | 201 | AAGVRCRLPGAYFFSFTLGKLPKTL | SVKLMKNRDEVOAMIYDDGASRRREMOSQSVML | 260 |
|----|-----|---|-----------------------------------|-----|
| | | | | |
| Db | 61 | AAGVRCRLPGAYFFSFTLGKLPKTL <th>SVKLMKNRDEVOAMIYDDGASRRREMOSQSVML</th> <th>120</th> | SVKLMKNRDEVOAMIYDDGASRRREMOSQSVML | 120 |

QY 261 ALRGDAVWLISHDGYGAYSNH 284
Db 121 ALRGDAVWLISHDGYGAYSNH 144
RESULT 10
ABG70383
ID ABG70383 standard; Protein; 221 AA.
XX
AC ABG70383;
XX
DT 05-NOV-2002 (first entry)
XX
DE Adipocyte complement-related C1q Tumour Necrosis Factor-like protein #1.
XX
KW Human; NOX; NOX-associated disorder; cardiomyopathy; atherosclerosis;
KW cell signal processing; metabolic pathway modulation; metabolic disorder;
KW obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
KW Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
KW haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
KW memory defect; infertility; congenital heart defect; hair growth;
KW pigmentation disorder; endocrine disorder; respiratory disease; health;
KW gastro-intestinal disease; reproductive; neurological disease;
KW bone marrow transplantation; endocrine disease; allergy; inflammation;
KW nephrological disorder; urinary system disorder; age-related disorder;
KW neuropsychiatric disorder; EGF-related protein; SCUBE1; TEN-M4;
KW adipocyte complement-related C1q tumour necrosis factor; out at first;
KW beta adrenergic receptor kinase; EphA6/enk-2; glucose transporter;
KW type Ia membrane sushi-containing domain; butyrophilin;
KW type Ia membrane-sushi domain containing.
XX
OS Homo sapiens.
XX
PN WO200257453-A2.
XX
PD 25-JUL-2002.
XX
PF 19-DEC-2001; 2001WO-US50331.
XX
PR 19-DEC-2000; 2000US-265704P.
PR 20-DEC-2000; 2000US-257314P.
PR 02-MAY-2001; 2001US-288153P.
PR 29-MAY-2001; 2001US-294075P.
PR 24-JUL-2001; 2001US-307506P.
PR 10-AUG-2001; 2001US-311590P.
PR 10-AUG-2001; 2001US-311613P.
PR 29-AUG-2001; 2001US-315617P.
PR 14-SEP-2001; 2001US-322358P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Gangolli EA, Patcurajan M, Vernet CAM, Malyankar UM, Kekuda R;
PI Stone DJ, Anderson D, Shimkets RA, Burgess CB, Zethusen BD, Liu X;
PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
XX WPI; 2002-590744/63.
DR N-PSDB; ABS52095.
XX
XX
PT Novel isolated NOX polypeptide useful for treating cardiomyopathy,
PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious
PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
PT cancer
XX
XX
PS Claim 1; Page 28; 318pp; English.
XX
CC The present invention relates to new NOX polypeptides. The invention is
CC useful for treating or preventing a NOX-associated disorder such as
CC cardiomyopathy or atherosclerosis, where the disorder is related to cell
CC signal processing and metabolic pathway modulation in a subject,
CC preferably human. The invention is also useful for treating metabolic
CC disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
CC disorders, haematopoietic disorders and various cancers. The molecules of

CC the invention are also useful for treating or preventing cirrhosis,
CC pancreatitis, learning and memory defects, infertility, congenital heart
CC defects, acne, hair growth, pigmentation disorders, reproductive health,
CC respiratory disease, gastro-intestinal diseases, reproductive health,
CC neurological diseases, bone marrow transplantation, endocrine diseases,
CC allergy and inflammation, nephrological disorders, urinary system
CC disorders, neuropsychiatric disorders and age-related disorders.
CC The present amino acid sequence represents a NOX protein of the
CC invention.
XX
SQ Sequence 221 AA;
XX
Query Match 43.2%; Score 707.5; DB 23; Length 221;
Best Local Similarity 49.5%; Pred. No. 4.8e-61;
Matches 153; Conservative 1; Mismatches 2; Indels 153; Gaps 1;
QY 5 PGPGSSSELSAFAAARTTPLEGTSSEMAVTPDKVYVNIIGDFDVATQFRCRVPGAYFFSF 64
Db 66 PGDSSAMRSASFSAARTTPLEGTSSEMAVTPDKVY----- 99
QY 65 TAGKAPHKSLSVMLVRNRDEVQALAFDEQRRPGARRAASQAMQLDYGDTVWLRLHGAP 124
Db 100 ----- 99
QY 125 HYALGAPGATFSGYLVYADADADAPARGPPAPPEPRSAFSAARTSLVGS DAGPRHQP 184
Db 100 ----- 99
QY 185 LAFDTEFVNIIGDFDAAGVFRCLPGAYFFSFTLGKLPRKTLVVKLMKNRDEVQAMITYD 244
Db 100 -----VNIGDFDAAGVFRCLPGAYFFSFTLGKLPRKTLVVKLMKNRDEVQAMITYD 152
QY 245 DGASRRREMOSQSVMLALRRGDAVWLISHDHDGYGAYSNHGKXITTFSGFLVYPDLAPAP 304
Db 153 DGASRRREMOSQSVMLALRRGDAVWLISHDHDGYGAYSNHGKXITTFSGFLVYPDLAPAP 212
QY 305 PGLGASELL 313
Db 213 PGLGASELL 221
RESULT 11
AAB61473
ID AAB61473 standard; protein; 134 AA.
XX
AC AAB61473;
XX
DT 04-APR-2001 (first entry)
XX
DE Monkey MANGO 245 C1q domain #2.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
XX Catarrhini sp..
OS
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases

PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Disclosure; Fig 28; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 134 AA;

Query Match 41.7%; Score 683; DB 22; Length 134;
Best Local Similarity 97.8%; Pred. No. 6.3e-59;
Matches 131; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 162 AFSARTSLVSGDAGPGRHQPLAFDTEFVNIGDPAAGVFRCLPGAYFSSFTLGK 221
ID ABB11578
ID ABB11578 standard; peptide; 127 AA.
XX
AC ABB11578;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human secreted protein homologue, SEQ ID NO:1948.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulneryary; antiulcer.
XX
OS Homo sapiens.
XX
PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03800.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-457740/49.

DR N-PSDB; ABA08822.
XX
PT Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX
PS Claim 20; Page 214-215; 1963pp; English.
XX
CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, and hence
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.
XX
SQ Sequence 127 AA;

Query Match 40.5%; Score 664; DB 22; Length 127;
Best Local Similarity 99.2%; Pred. No. 4.3e-57;
Matches 126; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 187 FDTFVNIGDPAAGVFRCLPGAYFSSFTLGKLPKRTLSVKLMKNRDEVQAMTYDDG 246
ID AAB61469
ID AAB61469 standard; protein; 126 AA.
XX
AC AAB61469;
XX
DT 04-APR-2001 (first entry)
XX
DR

DE Human MANGO 245 C1q domain #1.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Homo sapiens.
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Disclosure; Fig 26; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 126 AA;

Query Match 39.4%; Score 646; DB 22; Length 126;
Best Local Similarity 99.2%; Pred. No. 2.5e-55;
Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 15 AFSARTTPLEGTSEMAVTFDKVYVNIGDFDVATGQFRCRVPGAYFFSFTAGKAPHKSL 74
Db 1 AFSARTTPLEGTSEMAVTFDKVYVNIGDFDVATGQFRCRVPGAYFFSFTAGKAPHKSL 60

QY 75 SVMLVNRNDEVQALAFDEQRPGARRAASQASAMLQLDYGDTVWLRHLHGAPHYALGAPGAT 134
Db 61 SVMLVNRNDEVQALAFDEQRPGARRAASQASAMLQLDYGDTVWLRHLHGAPHYALGAPGAT 120

QY 135 FSGYLV 140
Db 121 FSGYLV 126

RESULT 14
AAB61489
ID AAB61489 standard; protein; 126 AA.
XX
AC AAB61489;
XX
DT 04-APR-2001 (first entry)
XX
DE Murine MANGO 245 C1q domain.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Mus musculus.
XX

PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Disclosure; Page 260-261; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 126 AA;

Query Match 39.0%; Score 639; DB 22; Length 126;
Best Local Similarity 96.8%; Pred. No. 1.2e-54;
Matches 122; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 15 AFSARTTPLEGTSEMAVTFDKVYVNIGDFDVATGQFRCRVPGAYFFSFTAGKAPHKSL 74
Db 1 AFSARTTPLEGTSEMAVTFDKVYVNIGDFDVATGQFRCRVPGAYFFSFTAGKAPHKSL 60

QY 75 SVMLVNRNDEVQALAFDEQRPGARRAASQASAMLQLDYGDTVWLRHLHGAPHYALGAPGAT 134
Db 61 SVMLVNRNDEVQALAFDEQRPGARRAASQASAMLQLDYGDTVWLRHLHGAPHYALGAPGAT 120

QY 135 FSGYLV 140
Db 121 FSGYLV 126

RESULT 15
AAB61477
ID AAB61477 standard; protein; 125 AA.
XX
AC AAB61477;
XX
DT 04-APR-2001 (first entry)
XX
DE Human MANGO 245 extracellular domain.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Homo sapiens.
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX

PA (MILL-) MILLENNIUM PHARM INC.
XX
XX Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
XX WPI; 2001-050127/06.
DR
XX
XX Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
XX
PS Disclosure; Page 246; 262pp; English.
XX
XX The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
XX
SQ Sequence 125 AA;

Query Match 38.7%; Score 635; DB 22; Length 125;
Best Local Similarity 100.0%; Pred.No. 2.9e-54;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGPTPGSSSELSAFAAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPAY 60
|||
Db 3 LGPTPGSSSELSAFAAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPAY 62
QY 61 FFSFTAGKAPHKSLVMLVRNRDEVQALAFDEQRRPGARRAASOSAMLQLDYGDTVWLRL 120
|||
Db 63 FFSFTAGKAPHKSLVMLVRNRDEVQALAFDEQRRPGARRAASOSAMLQLDYGDTVWLRL 122
QY 121 HGA 123
|||
Db 123 HGA 125

Search completed: January 12, 2004, 08:15:35
Job time : 48.1687 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:15:44 ; Search time 38.1098 Seconds
(without alignments)
1655.023 Million cell updates/sec

Title: US-10-085-167-2_COPY_17_329

Perfect score: 1639
Sequence: 1 LGFTPGGSSSELSAFAAAR.....LVYPLAPAPGLGASLL 313

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747907 seqs, 201509753 residues

Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|--------------------|--------------------|
| 1 | 1639 | 100.0 | 329 | 12 | US-10-085-167-2 | Sequence 2, Appli |
| 2 | 1633 | 99.6 | 329 | 12 | US-10-236-055A-14 | Sequence 14, Appli |
| 3 | 1514 | 92.4 | 326 | 12 | US-10-236-055A-16 | Sequence 16, Appli |
| 4 | 237.5 | 14.5 | 252 | 10 | US-09-893-737-52 | Sequence 52, Appli |
| 5 | 230 | 14.0 | 202 | 15 | US-10-091-458-43 | Sequence 43, Appli |
| 6 | 230 | 14.0 | 420 | 12 | US-10-236-055A-2 | Sequence 2, Appli |
| 7 | 230 | 14.0 | 746 | 10 | US-09-738-973-185 | Sequence 185, App |
| 8 | 230 | 14.0 | 746 | 10 | US-09-854-133-185 | Sequence 185, App |
| 9 | 230 | 14.0 | 746 | 15 | US-10-144-649A-185 | Sequence 185, App |
| 10 | 230 | 14.0 | 908 | 12 | US-10-309-422-10 | Sequence 10, Appli |
| 11 | 230 | 14.0 | 909 | 12 | US-10-309-422-22 | Sequence 12, Appli |
| 12 | 230 | 14.0 | 957 | 12 | US-10-309-422-14 | Sequence 14, Appli |
| 13 | 230 | 14.0 | 958 | 12 | US-10-309-422-26 | Sequence 26, Appli |
| 14 | 230 | 14.0 | 992 | 12 | US-10-309-422-18 | Sequence 18, Appli |
| 15 | 230 | 14.0 | 993 | 12 | US-10-309-422-30 | Sequence 30, Appli |

| | | | | | | |
|----|-------|------|------|----|-------------------|--------------------|
| 16 | 230 | 14.0 | 1042 | 12 | US-10-309-422-8 | Sequence 8, Appli |
| 17 | 230 | 14.0 | 1043 | 12 | US-10-309-422-20 | Sequence 20, Appli |
| 18 | 230 | 14.0 | 1091 | 12 | US-10-309-422-12 | Sequence 12, Appli |
| 19 | 230 | 14.0 | 1092 | 12 | US-10-309-422-24 | Sequence 24, Appli |
| 20 | 230 | 14.0 | 1126 | 12 | US-10-309-422-16 | Sequence 16, Appli |
| 21 | 230 | 14.0 | 1127 | 12 | US-10-309-422-28 | Sequence 28, Appli |
| 22 | 226 | 13.8 | 409 | 12 | US-10-236-055A-4 | Sequence 4, Appli |
| 23 | 220.5 | 13.5 | 281 | 10 | US-09-911-176B-44 | Sequence 44, Appli |
| 24 | 220.5 | 13.5 | 281 | 12 | US-10-360-186-44 | Sequence 44, Appli |
| 25 | 220.5 | 13.5 | 281 | 12 | US-10-392-531-44 | Sequence 44, Appli |
| 26 | 220.5 | 13.5 | 281 | 12 | US-10-392-706-44 | Sequence 44, Appli |
| 27 | 220.5 | 13.5 | 281 | 12 | US-10-385-015-5 | Sequence 5, Appli |
| 28 | 220.5 | 13.5 | 281 | 15 | US-10-180-762-44 | Sequence 44, Appli |
| 29 | 220.5 | 13.5 | 281 | 15 | US-10-241-258-44 | Sequence 44, Appli |
| 30 | 216 | 13.2 | 281 | 9 | US-09-768-826-46 | Sequence 46, Appli |
| 31 | 215 | 13.1 | 281 | 9 | US-09-765-205-24 | Sequence 24, Appli |
| 32 | 215 | 13.1 | 281 | 9 | US-09-866-028-78 | Sequence 78, Appli |
| 33 | 215 | 13.1 | 281 | 10 | US-09-944-449-78 | Sequence 78, Appli |
| 34 | 215 | 13.1 | 281 | 10 | US-09-944-457-78 | Sequence 78, Appli |
| 35 | 215 | 13.1 | 281 | 10 | US-09-944-862-78 | Sequence 78, Appli |
| 36 | 215 | 13.1 | 281 | 10 | US-09-945-587-78 | Sequence 78, Appli |
| 37 | 215 | 13.1 | 281 | 10 | US-09-945-015-78 | Sequence 78, Appli |
| 38 | 215 | 13.1 | 281 | 10 | US-09-944-396-78 | Sequence 78, Appli |
| 39 | 215 | 13.1 | 281 | 10 | US-09-944-097-78 | Sequence 78, Appli |
| 40 | 215 | 13.1 | 281 | 10 | US-09-944-432-78 | Sequence 78, Appli |
| 41 | 215 | 13.1 | 281 | 10 | US-09-943-762-78 | Sequence 78, Appli |
| 42 | 215 | 13.1 | 281 | 10 | US-09-944-654-78 | Sequence 78, Appli |
| 43 | 215 | 13.1 | 281 | 10 | US-09-943-851A-78 | Sequence 78, Appli |
| 44 | 215 | 13.1 | 281 | 10 | US-09-944-413-78 | Sequence 78, Appli |
| 45 | 215 | 13.1 | 281 | 10 | US-09-911-176B-2 | Sequence 2, Appli |

ALIGNMENTS

RESULT 1
US-10-085-167-2
; Sequence 2, Application US/10085167
; Publication No. US20030170781A1
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: SECRETED PROTEIN ZACRP4
; FILE REFERENCE: 99-29
; CURRENT APPLICATION NUMBER: US/10/085,167
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: 60/141,928
; PRIOR FILING DATE: 1999-07-01
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-085-167-2

| | | | | |
|-----------------------|-----------------|--|-----------|-------------|
| Query Match | 100.0%; | Score 1639; | DB 12; | Length 329; |
| Best Local Similarity | 100.0%; | Pred. No. 6.2e-149; | | |
| Matches 313; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| DB | 1 | LGFTPGGSSSELSAFAAARTTPLEGTSSEMAVTPDKVYVNIIGDPEVATGQFRCRVPGAY | 60 | |
| DB | 17 | LGFTPGGSSSELSAFAAARTTPLEGTSSEMAVTPDKVYVNIIGDPEVATGQFRCRVPGAY | 76 | |
| QY | 61 | FFSFTAGAKPHKSLVMLVRNRDEVOALAFDEQRRPGARRAASQASAMQLDYGDTVWLRL | 120 | |
| DB | 77 | FFSFTAGAKPHKSLVMLVRNRDEVOALAFDEQRRPGARRAASQASAMQLDYGDTVWLRL | 136 | |
| QY | 121 | HGAPHYALGAPGATFSGYLVVADADADAPARGPPAPPEPRSAFSAARTSLVGS DAGPGP | 180 | |
| DB | 137 | HGAPHYALGAPGATFSGYLVVADADADAPARGPPAPPEPRSAFSAARTSLVGS DAGPGP | 196 | |
| QY | 181 | RHQPLAFDTEFVNIGGDFDAAGVFRCLPGAYFFSFTLGKLRKTLVKLMKNRDEVQA | 240 | |

```
Db 197 RHQPLAFDTEFVNIGDPAAGVFRCLPGAYFFSFTLGKLPKTLGVKLMKNRDEVQA 256
QY 241 MIYDDGASRRREMOSQSVMLALRRGDAVWMLSHDHGYGAYSNHGKYITFSGFLVYPDLA 300
Db 257 MIYDDGASRRREMOSQSVMLALRRGDAVWMLSHDHGYGAYSNHGKYITFSGFLVYPDLA 316
QY 301 PAAPGGLGASELL 313
Db 317 PAAPGGLGASELL 329
```

RESULT 2

```
US-10-236-055A-14
; Sequence 14, Application US/10236055A
; Publication No. US20030134328A1
; GENERAL INFORMATION:
; APPLICANT: Basham, Beth E.
; APPLICANT: Forsythe, Ian
; APPLICANT: Gorman, Daniel M.
; APPLICANT: Mattson, Jeanine
; APPLICANT: Moshrefi, Mehrdad
; APPLICANT: Parham, Christi
; TITLE OF INVENTION: MAMMALIAN GENES; RELATED REAGENTS
; FILE REFERENCE: DX01343K
; CURRENT APPLICATION NUMBER: US/10/236,055A
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/317,988
; PRIOR FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-236-055A-14
```

Query Match 99.6%; Score 1633; DB 12; Length 329;
Best Local Similarity 99.7%; Pred. No. 2.3e-148;
Matches 312; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 LGPTPGGSSELSAFAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAY 60
Db 17 LGPTPGGSSELSAFAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAY 76
QY 61 FFSFTAGKAPHKSLSVMLVNRNDEVQALAFDEQRRPGARRAASQSAMQLDYGDTWMLRL 120
Db 77 FFSFTAGKAPHKSLSVMLVNRNDEVQALAFDEQRRPGARRAASQSAMQLDYGDTWMLRL 136
QY 121 HGAPHYALGAPGATFSGYLVYADADADAPARGPPAPPEPRSAFSAARTSLVGSADGPGP 180
Db 137 HGAPHYALGAPGATFSGYLVYADADADAPARGPPAPPEPRSAFSAARTSLVGSADGPGP 196
QY 181 RHQPLAFDTEFVNIGDPAAGVFRCLPGAYFFSFTLGKLPKTLGVKLMKNRDEVQA 240
Db 197 RHQPLAFDTEFVNIGDPAAGVFRCLPGAYFFSFTLGKLPKTLGVKLMKNRDEVQA 256
QY 241 MIYDDGASRRREMOSQSVMLALRRGDAVWMLSHDHGYGAYSNHGKYITFSGFLVYPDLA 300
Db 257 MIYDDGASRRREMOSQSVMLALRRGDAVWMLSHDHGYGAYSNHGKYITFSGFLVYPDLA 316
QY 301 PAAPGGLGASELL 313
Db 317 PAAPGGLGASELL 329
```

RESULT 3

```
US-10-236-055A-16
; Sequence 16, Application US/10236055A
; Publication No. US20030134328A1
; GENERAL INFORMATION:
; APPLICANT: Basham, Beth E.
; APPLICANT: Forsythe, Ian
```

```
; APPLICANT: Gorman, Daniel M.
; APPLICANT: Mattson, Jeanine
; APPLICANT: Moshrefi, Mehrdad
; APPLICANT: Parham, Christi
; TITLE OF INVENTION: MAMMALIAN GENES; RELATED REAGENTS
; FILE REFERENCE: DX01343K
; CURRENT APPLICATION NUMBER: US/10/236,055A
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/317,988
; PRIOR FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-236-055A-16
```

Query Match 92.4%; Score 1514; DB 12; Length 326;
Best Local Similarity 94.6%; Pred. No. 6.1e-137;
Matches 295; Conservative 2; Mismatches 13; Indels 2; Gaps 2;

```
QY 1 LGPTPGGSSELSAFAARTTPLEGTSEMAVTFDKVYVNIIGDFDVATGQFRCRVPGAY 60
Db 17 LGPA-GPGSSELSAFAARTTPLEGTSEMAVTFDKVYVNIIGDFDAATGRFRCRVPGAY 75
QY 61 FFSFTAGKAPHKSLSVMLVNRNDEVQALAFDEQRRPGARRAASQSAMQLDYGDTWMLRL 120
Db 76 FFSFTAGKAPHKSLSVMLVNRNDEVQALAFDEQRRPGARRAASQSAMQLDYGDTWMLRL 135
QY 121 HGAPHYALGAPGATFSGYLVYADADADAPARGPPAPPEPRSAFSAARTSLVGSADGPGP 180
Db 136 HGAPQYALGAPGATFSGYLVYADADADAPARG-PAAPPEPRSAFSAARTSLVGSADAPGP 194
QY 181 RHQPLAFDTEFVNIGDPAAGVFRCLPGAYFFSFTLGKLPKTLGVKLMKNRDEVQA 240
Db 195 RHRPLAFDTELVNIGDPAAGVFRCLPGAYFFSFTLGKLPKTLGVKLMKNRDEVQA 254
QY 241 MIYDDGASRRREMOSQSVMLALRRGDAVWMLSHDHGYGAYSNHGKYITFSGFLVYPDLA 300
Db 255 MIYDDGASRRREMOSQSVMLPLRRGDAVWMLSHDHGYGAYSNHGKYITFSGFLVYPDLA 314
QY 301 PAAPGGLGASEL 312
Db 315 AAGPPALKPPPEL 326
```

RESULT 4

```
US-09-893-737-52
; Sequence 52, Application US/09893737
; Patent No. US20020110855A1
; GENERAL INFORMATION:
; APPLICANT: Shepard, Paul O.
; APPLICANT: Presnell, Scott R.
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
; FILE REFERENCE: 00-41
; CURRENT APPLICATION NUMBER: US/09/893,737
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-893-737-52
```

Query Match 14.5%; Score 237.5; DB 10; Length 252;
Best Local Similarity 43.0%; Pred. No. 1.4e-14;
Matches 65; Conservative 12; Mismatches 61; Indels 13; Gaps 7;

QY 152 GPPAPPEPR--SAFSAARTSLVGSADGPPRHQPLAFDTEFVNIGDPAAGVFRCL 209

Db 106 GPPGACRRAYAAFSVGRREGHSSD-----HFOAVPFDTELVLNDGAFDLAAGRFLCTV 160
Qy 210 PGAYPFSFTLGKLPKTLVSKLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRRGDVW 269
Db 161 PGVYFLSLNVHTWNYKETYLHIMLR-RPAAVLYAQ-PSERSVMQAQSLMLLLAAGDAVW 218
Qy 270 --LLSHDHGCGAYSNHGK-YITFSGFLVYP 297
Db 219 VRMFORDRDN-AIYGEHGDLYITFSGHLVKP 248

RESULT 5

US-10-091-458-43
; Sequence 43, Application US/10091458
; Publication No. US20030068627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT206C1
; CURRENT APPLICATION NUMBER: US/10/091,458
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 09/764,900
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14


```

1 PRIOR APPLICATION NUMBER: 60/235, 836
2 PRIOR FILING DATE: 2000-09-27
3 PRIOR APPLICATION NUMBER: 60/230, 438
4 PRIOR FILING DATE: 2000-09-06
5 PRIOR APPLICATION NUMBER: 60/215, 135
6 PRIOR FILING DATE: 2000-06-30
7 PRIOR APPLICATION NUMBER: 60/225, 266
8 PRIOR FILING DATE: 2000-08-14
9 PRIOR APPLICATION NUMBER: 60/249, 218
10 PRIOR FILING DATE: 2000-11-17
11 PRIOR APPLICATION NUMBER: 60/249, 208
12 PRIOR FILING DATE: 2000-11-17
13 PRIOR APPLICATION NUMBER: 60/249, 213
14 PRIOR FILING DATE: 2000-11-17
15 PRIOR APPLICATION NUMBER: 60/249, 212
16 PRIOR FILING DATE: 2000-11-17
17 PRIOR APPLICATION NUMBER: 60/249, 207
18 PRIOR FILING DATE: 2000-11-17
19 PRIOR APPLICATION NUMBER: 60/249, 245
20 PRIOR FILING DATE: 2000-11-17
21 PRIOR APPLICATION NUMBER: 60/249, 244
22 PRIOR FILING DATE: 2000-11-17
23 PRIOR APPLICATION NUMBER: 60/249, 217
24 PRIOR FILING DATE: 2000-11-17
25 PRIOR APPLICATION NUMBER: 60/249, 211
26 PRIOR FILING DATE: 2000-11-17
27 PRIOR APPLICATION NUMBER: 60/249, 215
28 PRIOR FILING DATE: 2000-11-17
29 PRIOR APPLICATION NUMBER: 60/249, 264
30 PRIOR FILING DATE: 2000-11-17
31 PRIOR APPLICATION NUMBER: 60/249, 214
32 PRIOR FILING DATE: 2000-11-17
33 PRIOR APPLICATION NUMBER: 60/249, 297
34 PRIOR FILING DATE: 2000-11-17
35 PRIOR APPLICATION NUMBER: 60/232, 400
36 PRIOR FILING DATE: 2000-09-14
37 PRIOR APPLICATION NUMBER: 60/231, 242
38 PRIOR FILING DATE: 2000-09-08
39 PRIOR APPLICATION NUMBER: 60/232, 081
40 PRIOR FILING DATE: 2000-09-08
41 PRIOR APPLICATION NUMBER: 60/232, 080
42 PRIOR FILING DATE: 2000-09-08
43 PRIOR APPLICATION NUMBER: 60/231, 414
44 PRIOR FILING DATE: 2000-09-08
45 PRIOR APPLICATION NUMBER: 60/231, 244
46 PRIOR FILING DATE: 2000-09-08
47 PRIOR APPLICATION NUMBER: 60/233, 064
48 PRIOR FILING DATE: 2000-09-14
49 PRIOR APPLICATION NUMBER: 60/233, 063
50 PRIOR FILING DATE: 2000-09-14
51 PRIOR APPLICATION NUMBER: 60/232, 397
52 PRIOR FILING DATE: 2000-09-14
53 PRIOR APPLICATION NUMBER: 60/232, 399
54 PRIOR FILING DATE: 2000-09-14
55 PRIOR APPLICATION NUMBER: 60/232, 401
56 PRIOR FILING DATE: 2000-09-14
57 PRIOR APPLICATION NUMBER: 60/241, 808
58 PRIOR FILING DATE: 2000-10-20
59 PRIOR APPLICATION NUMBER: 60/241, 826
60 PRIOR FILING DATE: 2000-10-20
61 PRIOR APPLICATION NUMBER: 60/241, 786
62 PRIOR FILING DATE: 2000-10-20
63 PRIOR APPLICATION NUMBER: 60/241, 221
64 PRIOR FILING DATE: 2000-10-20
65 PRIOR APPLICATION NUMBER: 60/246, 475
66 PRIOR FILING DATE: 2000-11-08
67 PRIOR APPLICATION NUMBER: 60/231, 243
68 PRIOR FILING DATE: 2000-09-08

```

```

QY      147 DAPARGP-----PAPPEPSAFSAARTSLVSGDAGPGPRHQPLAEDTEFVNIGG 196
      ||| ||| : ||| ||| ||| : ||| ||| ||| : |||
Db      49 DVEVTNPAATILPVHVYPLPQGMKVAFSAART-----SNLAPGTLDPIVEFDLLNNLGE 103
      ||| ||| : ||| ||| ||| : ||| ||| ||| : |||
QY      197 DFDAAAGVFRCLPGRAYFFSFTLGKLP-RKTL SVKLMKNRDEVQAMIYDDGASRRREMQS 255
      ||| ||| : ||| ||| ||| : ||| ||| ||| : |||
Db      104 TFDLQLGFRNCPVNGTYVFI FHM LKLA VNVPLVNLMKNEEVLVSAYANDGAP-DHETAS 162
      ||| ||| : ||| ||| ||| : ||| ||| ||| : |||
QY      256 QSVMLALRRGDVAVMLLSHDHDGYGA-YSNHGKIYTFSGFLVYPD 298
      : ||| : ||| : ||| ||| : ||| ||| ||| : |||
Db      163 NHAIIQLFQGDQIWLRLH---RCAIYGSSWKIYSTFSGYLLLYQD 202

```

RESULT 6
US-10-236-055A-2

```

; Sequence 2, Application US/10236055A
; Publication No. US20030134328A1
; GENERAL INFORMATION:
; APPLICANT: Basham, Beth E.
; APPLICANT: Forsythe, Ian
; APPLICANT: Gorman, Daniel M.
; APPLICANT: Mattson, Jeanine
; APPLICANT: Moshrefi, Mehrdad
; APPLICANT: Parham, Christi
; TITLE OF INVENTION: MAMMALIAN GENES; RELATED REAGENTS
; FILE REFERENCE: DX01343X
; CURRENT APPLICATION NUMBER: US/10/236,055A
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/317,988
; PRIOR FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-236-055A-2

```

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 14.0%; | Score 230; | DB 12; | Length 420; |
| Best Local Similarity | 38.4%; | Pred. No. 1.5e-13; | | |
| Matches 63; | Conservative 16; | Mismatches 63; | Indels 22; | Gaps 6; |

QY 147 DAPARGP-----PAPPEPSAESAARTSLVSGDAGPGPRHQPLAEDTEFVNIGG 196
| | | | | : | | | | | : | | | | | : | |
Db 267 DVEPTNPATILPVHVYPLPQCMRVAFSAART-----SNLAPGTLDPIVEDLLNNLGE 321
| | | | | : | | | | | : | | | | | : | |
QY 197 DEDAAGVFERCRLPGAYFFSFITGKLP-RKTLSTVKLMKNRDEVQAMIYDDGASRRREMOS 255
| | | | | : | | | | | : | | | | | : | | | | |
Db 322 TFDLQIGRFNCVPVNGTYVFI FHMCLKLAVNVPLVNLMKNEEVLVSAYANDGAP-DHETAS 380
| | | | | : | | | | | : | | | | | : | | | | |
QY 256 QSVMLALRRGDAVWLLSHDHGTYGA-YSNHGKIITFSGLVYPD 298
: | | : | | : | | : | | : | | : | | : | | : | |
Db 381 NHAITOLFQGDQIWLRLH---RGAIYGSWSWKYSTFSGYLLLYQD 420

RESULT 7
US-09-738-973-185

```

: Sequence 185, Application US/09738973
: Patent No. US20020110563A1
: GENERAL INFORMATION:
: APPLICANT: Reed, Steven G.
: APPLICANT: Henderson, Robert A.
: APPLICANT: Lodes, Michael J.
: APPLICANT: Fling, Steven P.
: APPLICANT: Mohamath, Raodoh
: APPLICANT: Algate, Paul A.
: APPLICANT: Secrist, Heather
: APPLICANT: Indirias, Carol Yoseph
: APPLICANT: Benson, Darin R.
: APPLICANT: Elliot, Mark
: APPLICANT: Mannion, Jane
: APPLICANT: Kalos, Michael D.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR

```

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 14.0%; | Score 230; | DB 15; | Length 202; |
| Best Local Similarity | 38.4%; | Pred. No. 5.5e-14; | | |
| Matches 63; | Conservative 16; | Mismatches 63; | Indels 22; | Gaps 6; |

```

; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 185
; LENGTH: 746
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-738-973-185

Query Match          14.0%; Score 230; DB 10; Length 746;
Best Local Similarity 38.4%; Pred. No. 3.2e-13;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

QY 147 DAPARGP-----PAPPEPSAFSAARTRSLVSGDAGPGRHQPLAFDTEFVNIGG 196
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 593 DVPVTNPAATILPVHVYPLPQQMRVAFSAART-----SNLAPGTLDPQPIVFDLLNNLGE 647

QY 197 DFDAAGVFRCLPGAYFFSFTLGKL-P-RKTLISVKLMKNRDEVQAMITYDDGASRRREMOS 255
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 648 TFDLQIGRFNCPVNGTYVFI FHM LKLA VNVPLVY NLMKN EEVLSA YANDGAP-DHETAS 706

QY 256 QSVMLALRRGDAVWL LSHDHDGYGA-YSNHGKYITFSGFLVYPD 298
   :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||
DB 707 NHAIIQLFGQDIWLR L H---RGA IYGS SWKYSTFSGYLLYQD 746

RESULT 8
US-09-854-133-185
; Sequence 185, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 185
; LENGTH: 746
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-854-133-185

Query Match          14.0%; Score 230; DB 10; Length 746;
Best Local Similarity 38.4%; Pred. No. 3.2e-13;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

QY 147 DAPARGP-----PAPPEPSAFSAARTRSLVSGDAGPGRHQPLAFDTEFVNIGG 196
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 593 DVPVTNPAATILPVHVYPLPQQMRVAFSAART-----SNLAPGTLDPQPIVFDLLNNLGE 647

QY 197 DFDAAGVFRCLPGAYFFSFTLGKL-P-RKTLISVKLMKNRDEVQAMITYDDGASRRREMOS 255
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 648 TFDLQIGRFNCPVNGTYVFI FHM LKLA VNVPLVY NLMKN EEVLSA YANDGAP-DHETAS 706

QY 256 QSVMLALRRGDAVWL LSHDHDGYGA-YSNHGKYITFSGFLVYPD 298
   :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||
DB 707 NHAIIQLFGQDIWLR L H---RGA IYGS SWKYSTFSGYLLYQD 746

RESULT 9
US-10-144-649A-185
; Sequence 185, Application US/10144649A
; Publication No. US20030118599A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tonglong
; APPLICANT: Fan, Liqun
; APPLICANT: Algate, Paul A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C11
; CURRENT APPLICATION NUMBER: US/10/144,649A
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 749
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 185
; LENGTH: 746
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-144-649A-185

Query Match          14.0%; Score 230; DB 15; Length 746;
Best Local Similarity 38.4%; Pred. No. 3.2e-13;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

QY 147 DAPARGP-----PAPPEPSAFSAARTRSLVSGDAGPGRHQPLAFDTEFVNIGG 196
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 593 DVPVTNPAATILPVHVYPLPQQMRVAFSAART-----SNLAPGTLDPQPIVFDLLNNLGE 647

QY 197 DFDAAGVFRCLPGAYFFSFTLGKL-P-RKTLISVKLMKNRDEVQAMITYDDGASRRREMOS 255
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 648 TFDLQIGRFNCPVNGTYVFI FHM LKLA VNVPLVY NLMKN EEVLSA YANDGAP-DHETAS 706

QY 256 QSVMLALRRGDAVWL LSHDHDGYGA-YSNHGKYITFSGFLVYPD 298
   :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||
DB 707 NHAIIQLFGQDIWLR L H---RGA IYGS SWKYSTFSGYLLYQD 746

RESULT 10
US-10-309-422-10
; Sequence 10, Application US/10309422
; Publication No. US20030139587A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: NO. US20030139587A1e1 Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: LEX-0142-USA
; CURRENT APPLICATION NUMBER: US/10/309,422
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: US/09/798,771
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,557
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 908
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-309-422-10

Query Match          14.0%; Score 230; DB 12; Length 908;
Best Local Similarity 38.4%; Pred. No. 4.2e-13;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

QY 147 DAPARGP-----PAPPEPSAFSAARTRSLVSGDAGPGRHQPLAFDTEFVNIGG 196
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 755 DVPVTNPAATILPVHVYPLPQQMRVAFSAART-----SNLAPGTLDPQPIVFDLLNNLGE 809

QY 197 DFDAAGVFRCLPGAYFFSFTLGKL-P-RKTLISVKLMKNRDEVQAMITYDDGASRRREMOS 255
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 810 TFDLQIGRFNCPVNGTYVFI FHM LKLA VNVPLVY NLMKN EEVLSA YANDGAP-DHETAS 868

QY 256 QSVMLALRRGDAVWL LSHDHDGYGA-YSNHGKYITFSGFLVYPD 298
```


LENGTH: 992
TYPE: PRT
ORGANISM: homo sapiens
US-10-309-422-18

Query Match 14.0%; Score 230; DB 12; Length 992;
Best Local Similarity 38.4%; Pred. No. 4.7e-13;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

QY 147 DAPARGP-----PAPPEPSAFSAARTSLVGS DAGPGRHQPLAFDTEFVNIGG 196
DB 839 DVPVTNPATILPVHYVPLPQOMRVAFSAART-----SNLAPGTLDDPIVFDLLLNLTGE 893
QY 197 DFDAAGVFRCLPGAYFFSFTLGKLP-RKTLGVKLMKNRDEVQAMTYDDGASRRREMOS 255
DB 894 TFDLQIGRFNCPVNGTYVFI FHM LKLA VNPPLYVNL MKN E EVLSA YANDGAP-DHETAS 952
QY 256 QSVMLALRRGDVWLLSHDHDGYGA-YSNNGKYITFSGFLVYPD 298
DB 953 NHA ILQLFQGDQIWLR LH---RGAIY GSSWKYSTFSGYLLYOD 992

RESULT 15
US-10-309-422-30
; Sequence 30, Application US/10309422
; Publication No. US20030139587A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20030139587A1e1 Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: LEX-0142-USA
; CURRENT APPLICATION NUMBER: US/10/309,422
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: US/09/798,771
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,557
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 30
; LENGTH: 993
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-309-422-30

Query Match 14.0%; Score 230; DB 12; Length 993;
Best Local Similarity 38.4%; Pred. No. 4.7e-13;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

QY 147 DAPARGP-----PAPPEPSAFSAARTSLVGS DAGPGRHQPLAFDTEFVNIGG 196
DB 840 DVPVTNPATILPVHYVPLPQOMRVAFSAART-----SNLAPGTLDDPIVFDLLLNLTGE 894
QY 197 DFDAAGVFRCLPGAYFFSFTLGKLP-RKTLGVKLMKNRDEVQAMTYDDGASRRREMOS 255
DB 895 TFDLQIGRFNCPVNGTYVFI FHM LKLA VNPPLYVNL MKN E EVLSA YANDGAP-DHETAS 953
QY 256 QSVMLALRRGDVWLLSHDHDGYGA-YSNNGKYITFSGFLVYPD 298
DB 954 NHA ILQLFQGDQIWLR LH---RGAIY GSSWKYSTFSGYLLYOD 993

Search completed: January 12, 2004, 08:33:36
Job time : 39.1098 secs

THIS PAGE BLANK (USPTO)


```

; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/053,154
; EARLIER FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 44
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-118-408-44

```

| | | | | |
|-----------------------|------------------|--------------------|-----------|-------------|
| Query Match | 13.5%; | Score 220.5; | DB 3; | Length 281; |
| Best Local Similarity | 39.2%; | Pred. No. 1.8e-14; | | |
| Matches 56; | Conservative 20; | Mismatches 58; | Indels 9; | Gaps 4; |

```

QY      161 SAFSAARTRSLVGSADGPGPRHOPIAFDTEFVNIGDFFDAAGVFRCLPGAVFFSFTLG  220
      :||| |::| :| :|:|||||:| :| :|:||||| :|
Db      146 AAFSVGRKKALHSND----YFQPVVFDETFVNLKHFNMFTGKFCYVPGIYFFSLNVH  200
QY      221 KLPRKTLVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRRGDVW--LLSHDHGX  278
      :| :|:||||:| :| :|:||||:| :| :|:||||:| :|
Db      201 TWNQKETYLHIMKNEEV-VILYAQ-VSDRSIMOSQSLMELRDEDEVWVRLFKGERENA  258
QY      279 GAYSNHGKIYTFSGFLVYPDLAP 301
      |||||:|:| |
Db      259 IFSDEFDTYITFSGYLVKPPASEP 281

```

RESULT 3
US-09-50

US-09-506-855-44
; Sequence 44, Application US/09506855

```

; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Laesser, Gerald W.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: INHIBITORS FOR USE IN HEMOSTASIS AND
; TITLE OF INVENTION: IMMUNE FUNCTION
; FILE REFERENCE: 99-12
; CURRENT APPLICATION NUMBER: US/09/506,855
; CURRENT FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 44
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-506-855-44

```

| | | | | |
|-----------------------|------------------|--------------------|-----------|-------------|
| Query Match | 13.5%; | Score 220.5; | DB 4; | Length 281; |
| Best Local Similarity | 39.2%; | Pred. No. 1.8e-14; | | |
| Matches 56; | Conservative 20; | Mismatches 58; | Indels 9; | Gaps 4; |

[illegible]

RESULT 4

US-09-911-176B-44

; Sequence 44, Application US/09911176B

Patent No. 6518403

APPLICANT: Sheppard, Paul O.

1 TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOG
2
3 FILE REFERENCE: 97-30D1
4

```

; FILE REFERENCE: 97-3001
; CURRENT APPLICATION NUMBER: US/09/911,176B
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: 09/118,408
; PRIOR FILING DATE: 1998-07-17
; PRIOR APPLICATION NUMBER: 60/053,154
; PRIOR FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 44
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-911-176B-44

```

| | | | | |
|-----------------------|------------------|--------------------|-----------|-------------|
| Query Match | 13.5%; | Score 220.5; | DB 4; | Length 281; |
| Best Local Similarity | 39.2%; | Pred. No. 1.8e-14; | | |
| Matches 56; | Conservative 20; | Mismatches 58; | Indels 9; | Gaps 4; |

| | | | |
|----|-----|--|-----------------------|
| QY | 161 | SAFSAARTSLVGSADAGPGRHQPLAFDTEFVNIGGDFDAAGVRCRLPGAYFFSFTLG | 220 |
| | : | : : : | : : : : |
| Db | 146 | AASFVGRRKALHSND-----YFQPVEVDTEFVNLYKHFNMFGTKFYCYVPGIYFFSLNVH | 200 |
| QY | 221 | KLPKRTLSVKLMKNRDEVQAMITDDGASRRREMOSQSVMALARRGDAW--LLSHDHDX | 278 |
| | : | : : : : : | : : : |
| Db | 201 | TWNQKETYLHIKMNEEV-VILYAQ-VSDRSIMOSQLMELREDEWVRLLFKGERENA | 258 |
| QY | 279 | GAYSNHGYITFSGFVYPDLAP | 301 |
| | : | : : | |
| Db | 259 | IFSDEFDTYITFSGYLVKRPASEP | 281 |

RESULT 5

US-09-619-740-44

Sequence 44, Application US/09619740

Patent No. 6544946

```

; GENERAL INFORMATION:
;
APPLICANT: Shennard, Paul O.

```

APPLICANT: Hassler, Gerald W.

APPLICANT: Bishop, Paul D.

TITLE OF INVENTION: INHIBITORS OF

FILE REFERENCE: 99-12C3

CURRENT APPLICATION NUMBER: US/09/619

CURRENT FILING DATE: 2000-07-19
;
PRIORITY FILING DATE: 09/25/99
PRIORITY FILING NUMBER: 604

PRIOR APPLICATION NUMBER: 09/253,604
PRIOR FILING DATE: 1999-02-19

PRIOR FILING DATE: 1999-02-19
PRIOR APPLICATION NUMBER: 09/444,794

PRIOR FILING DATE: 1999-11-22

PRIOR APPLICATION NUMBER: 09/506,855

PRIOR FILING DATE: 2000-02-17

; NUMBER OF SEQ ID NOS: 55

; SOFTWARE: F

; SEQ ID NO 44
LENGTH: 283

LENGTH: 281
TYPE: PRT

ORGANISM: *Mus musculus*

US-09-619-740-44

| | | | | |
|--------------------------|--------|--------------------|-----------|-------------|
| Query Match | 13.5% | Score 220.5; | DB 4; | Length 281; |
| Best Local Similarity | 39.2%; | Pred. No. 1.8e-14; | | |
| Matches 56; Conservative | 20; | Mismatches 58; | Indels 9; | Gaps 4; |

```

QY      .      161 SAFSARTRSLVGSADGPPRRHQPLAFDTEFVNIGGDFDAAGVFCRCLPGAYFFSFTLG 220
      :||| |::| :| :|:|||||: | :| :|:||||| :
Db      146 AAFSVGRKKALHSND----YQDPVVFDTFVNLYKHFMNMTGKFYCYVPGIYFFSLNVH 200

```

```
QY      221 KLPRKTLVKLMKNRDEVQAITYDDGASRRREMOSQSVMALLRGDAVW--LLSHDHG DY 278
       :| :||| :|| :| | ||||| :| | | | :|
Db     201 TWNOKETYLHIMKNEEV-VILTAO-VSDRSIMOSQLMELREDEVTVRLLFKGERENA 258
```

QY 279 GAYSNHGKYITFSGFLVYPDLAP 301


```
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-911-176B-2
```

```
Query Match      13.1%; Score 215; DB 4; Length 281;
Best Local Similarity 30.2%; Pred. No. 6.6e-14;
Matches 77; Conservative 26; Mismatches 86; Indels 66; Gaps 11;
```

```
OY 53 RCRVPGAYFFSFTAGKAPHKSLVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDY 112
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 75 RCCDPGTSMPATA--VPQINITILKGEKDR-----GDRG-----LQGY 113

OY 113 GDT--VWLRHLGAP--HYALGAPGATFSGYLVADADADAPARGPAPPEPRSAFSAAR 167
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 114 GKTGSAGARGHTGPKGQKSGMGAPEGCKSH--YA-----AFSVGR 152

OY 168 TRSLVGSADAGPGRH-----QPLAFDTEFVNIGDFAAAGVFRCLPGAYFFSFTLGKL 222
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 153 KK-----PMHSNHYQTIVIFDTEFVNLYDHFMFTGKFYCYVPGLYFFSLNVHTW 202

OY 223 PRKTLVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRRGDVWMLSHDHGYGAYS 282
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 203 NQKETYLHMKNEEEVILFAQVG--DRSIMQSQSLMLELREODQVWVRLYKGERENAIF 260

OY 283 NH--GKYITFSGFLV 295
    :|||||:|||||:
Db 261 SEELDTYITFSGYLV 275
```

```
RESULT 10
US-09-619-740-2
; Sequence 2, Application US/09619740
; Patent No. 6544946
; GENERAL INFORMATION:
```

```
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Lasser, Gerald W.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: INHIBITORS FOR USE IN HEMOSTASIS AND IMMUNE FUNCTION
; FILE REFERENCE: 99-12C3
; CURRENT APPLICATION NUMBER: US/09/619,740
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/253,604
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/444,794
; PRIOR FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: 09/506,855
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-619-740-2
```

```
Query Match      13.1%; Score 215; DB 4; Length 281;
Best Local Similarity 30.2%; Pred. No. 6.6e-14;
Matches 77; Conservative 26; Mismatches 86; Indels 66; Gaps 11;
```

```
OY 53 RCRVPGAYFFSFTAGKAPHKSLVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDY 112
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 75 RCCDPGTSMPATA--VPQINITILKGEKDR-----GDRG-----LQGY 113

OY 113 GDT--VWLRHLGAP--HYALGAPGATFSGYLVADADADAPARGPAPPEPRSAFSAAR 167
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 114 GKTGSAGARGHTGPKGQKSGMGAPEGCKSH--YA-----AFSVGR 152

OY 168 TRSLVGSADAGPGRH-----QPLAFDTEFVNIGDFAAAGVFRCLPGAYFFSFTLGKL 222
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db 153 KK-----PMHSNHYQTIVIFDTEFVNLYDHFMFTGKFYCYVPGLYFFSLNVHTW 202

OY 223 PRKTLVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRRGDVWMLSHDHGYGAYS 282
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 203 NQKETYLHMKNEEEVILFAQVG--DRSIMQSQSLMLELREODQVWVRLYKGERENAIF 260

OY 283 NH--GKYITFSGFLV 295
    :|||||:|||||:
Db 261 SEELDTYITFSGYLV 275
```

```
RESULT 11
```

```
US-09-506-852-2
; Sequence 2, Application US/09506852
; Patent No. 6566499
; GENERAL INFORMATION:
```

```
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
; FILE REFERENCE: 97-30
; CURRENT APPLICATION NUMBER: US/09/506,852
; CURRENT FILING DATE: 2000-02-17
; EARLIER APPLICATION NUMBER: 60/053,154
; EARLIER FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-506-852-2
```

```
Query Match      13.1%; Score 215; DB 4; Length 281;
Best Local Similarity 30.2%; Pred. No. 6.6e-14;
Matches 77; Conservative 26; Mismatches 86; Indels 66; Gaps 11;
```

```
OY 53 RCRVPGAYFFSFTAGKAPHKSLVMLVRNDEVQALAFDEQRRPGARRAASQSAMLQLDY 112
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 75 RCCDPGTSMPATA--VPQINITILKGEKDR-----GDRG-----LQGY 113

OY 113 GDT--VWLRHLGAP--HYALGAPGATFSGYLVADADADAPARGPAPPEPRSAFSAAR 167
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 114 GKTGSAGARGHTGPKGQKSGMGAPEGCKSH--YA-----AFSVGR 152

OY 168 TRSLVGSADAGPGRH-----QPLAFDTEFVNIGDFAAAGVFRCLPGAYFFSFTLGKL 222
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 153 KK-----PMHSNHYQTIVIFDTEFVNLYDHFMFTGKFYCYVPGLYFFSLNVHTW 202

OY 223 PRKTLVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRRGDVWMLSHDHGYGAYS 282
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 203 NQKETYLHMKNEEEVILFAQVG--DRSIMQSQSLMLELREODQVWVRLYKGERENAIF 260

OY 283 NH--GKYITFSGFLV 295
    :|||||:|||||:
Db 261 SEELDTYITFSGYLV 275
```

```
RESULT 12
```

```
US-09-336-536-4
; Sequence 4, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
```

```
; APPLICANT: leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 228
; TYPE: PRT
```

```

; ORGANISM: Homo sapiens
US-09-336-536-4

Query Match      13.0%; Score 213; DB 4; Length 228;
Best Local Similarity 34.9%; Pred. No. 7.9e-14;
Matches 66; Conservative 18; Mismatches 83; Indels 22; Gaps 6;

QY 122 GAPHYALGAPGATFSG---YLVIYADAD---ADAPARGPPAPP-----EPRSAFSAARTR 169
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 42 GAP----GAPGKEGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSASFSAKRSE 97

QY 170 SLVGS DAGPGRHQPLAFDTEFVNIGDFDAAGVFRCLPGAYFFSFTLGKLPKRTL SV 229
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 98 SRV-----PPPSDAPLPFDRVLVNEQGHYDAVTGKTCQVPGVYYFA-VHATVYRASLQF 151

QY 230 KLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRRGDAVWL LSHDHG YGAYSNHGKYIT 289
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152 DLVKNGESIASFFQFFGWP KPASLSGAMVRLPEPDQVWVQVGVDYIGIYASIKTDST 211

QY 290 FSGFLVYPD 298
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 212 FSGFLVYSD 220

RESULT 13
US-09-140-804-2
; Sequence 2, Application US/09140804
; Patent No. 6197930
; GENERAL INFORMATION:
; APPLICANT: Shepard, Paul O.
; APPLICANT: Humes, Jacqueline M.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
; FILE REFERENCE: 97-49
; CURRENT APPLICATION NUMBER: US/09/140,804
; CURRENT FILING DATE: 1998-08-26
; EARLIER APPLICATION NUMBER: 60/056,983
; EARLIER FILING DATE: 1997-08-26
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-140-804-2

Query Match      13.0%; Score 213; DB 3; Length 243;
Best Local Similarity 34.9%; Pred. No. 8.6e-14;
Matches 66; Conservative 18; Mismatches 83; Indels 22; Gaps 6;

QY 122 GAPHYALGAPGATFSG---YLVIYADAD---ADAPARGPPAPP-----EPRSAFSAARTR 169
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 57 GAP----GAPGKEGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSASFSAKRSE 112

QY 170 SLVGS DAGPGRHQPLAFDTEFVNIGDFDAAGVFRCLPGAYFFSFTLGKLPKRTL SV 229
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 SRV-----PPPSDAPLPFDRVLVNEQGHYDAVTGKTCQVPGVYYFA-VHATVYRASLQF 166

QY 230 KLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRRGDAVWL LSHDHG YGAYSNHGKYIT 289
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 167 DLVKNGESIASFFQFFGWP KPASLSGAMVRLPEPDQVWVQVGVDYIGIYASIKTDST 226

QY 290 FSGFLVYPD 298
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 227 FSGFLVYSD 235

RESULT 14
US-09-336-536-3
; Sequence 3, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
```

```

; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-336-536-3

Query Match      13.0%; Score 213; DB 4; Length 243;
Best Local Similarity 34.9%; Pred. No. 8.6e-14;
Matches 66; Conservative 18; Mismatches 83; Indels 22; Gaps 6;

QY 122 GAPHYALGAPGATFSG---YLVIYADAD---ADAPARGPPAPP-----EPRSAFSAARTR 169
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 57 GAP----GAPGKEGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSASFSAKRSE 112

QY 170 SLVGS DAGPGRHQPLAFDTEFVNIGDFDAAGVFRCLPGAYFFSFTLGKLPKRTL SV 229
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 SRV-----PPPSDAPLPFDRVLVNEQGHYDAVTGKTCQVPGVYYFA-VHATVYRASLQF 166

QY 230 KLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRRGDAVWL LSHDHG YGAYSNHGKYIT 289
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 167 DLVKNGESIASFFQFFGWP KPASLSGAMVRLPEPDQVWVQVGVDYIGIYASIKTDST 226

QY 290 FSGFLVYPD 298
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 227 FSGFLVYSD 235
```

```

RESULT 15
US-09-686-838B-2
; Sequence 2, Application US/09686838B
; Patent No. 6482612
; GENERAL INFORMATION:
; APPLICANT: Shepard, Paul O.
; APPLICANT: Humes, Jacqueline M.
; TITLE OF INVENTION: Adipocyte-Specific Protein Homologs
; FILE REFERENCE: 97-49D1
; CURRENT APPLICATION NUMBER: US/09/686,838B
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/140,804
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: US 60/056,983
; PRIOR FILING DATE: 1997-08-26
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-686-838B-2

Query Match      13.0%; Score 213; DB 4; Length 243;
Best Local Similarity 34.9%; Pred. No. 8.6e-14;
Matches 66; Conservative 18; Mismatches 83; Indels 22; Gaps 6;

QY 122 GAPHYALGAPGATFSG---YLVIYADAD---ADAPARGPPAPP-----EPRSAFSAARTR 169
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 57 GAP----GAPGKEGEGGRPGLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSASFSAKRSE 112

QY 170 SLVGS DAGPGRHQPLAFDTEFVNIGDFDAAGVFRCLPGAYFFSFTLGKLPKRTL SV 229
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 SRV-----PPPSDAPLPFDRVLVNEQGHYDAVTGKTCQVPGVYYFA-VHATVYRASLQF 166

QY 230 KLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRRGDAVWL LSHDHG YGAYSNHGKYIT 289
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 167 DLVKNGESIASFFQFFGWP KPASLSGAMVRLPEPDQVWVQVGVDYIGIYASIKTDST 226
```

| | | | |
|----|-----|-----------|-----|
| Qy | 290 | FSGFLVYPD | 298 |
| | | | |
| Db | 227 | FSGFLVYSD | 235 |

Search completed: January 12, 2004, 08:20:04
Job time : 16.9311 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:09 ; Search time 9.84032 Seconds
(without alignments)
1661.397 Million cell updates/sec

Title: US-10-085-167-2_COPY_160_358
Perfect score: 894
Sequence: 1 ADADAPARGPPAPPEPRSAF.....LVYPDLAPAAPGLGASELL 170

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Query Length | DB ID | Description |
|------------|-------|-------------|--------------|-------|---------------------------|
| 1 | 208.5 | 23.3 | 219 | 2 | T14782 hypothetical prote |
| 2 | 182 | 20.4 | 253 | 1 | CIHUOB complement subcomp |
| 3 | 177.5 | 19.9 | 215 | 2 | B48150 hibernation-relate |
| 4 | 175 | 19.6 | 244 | 2 | JC4708 gelatin-binding 28 |
| 5 | 170 | 19.0 | 245 | 1 | CIHUOC complement subcomp |
| 6 | 166.5 | 18.6 | 674 | 2 | S23297 collagen alpha 1(X |
| 7 | 165.5 | 18.5 | 253 | 2 | S49158 complement protein |
| 8 | 165.5 | 18.5 | 744 | 1 | S23298 collagen alpha 1(V |
| 9 | 164 | 18.3 | 680 | 1 | CGHUID collagen alpha 1(X |
| 10 | 164 | 18.3 | 680 | 2 | S31216 collagen alpha 1(X |
| 11 | 163.5 | 18.3 | 744 | 1 | A34246 collagen alpha 1(X |
| 12 | 163 | 18.2 | 224 | 2 | A60032 cerebellin-like g1 |
| 13 | 163 | 18.2 | 674 | 2 | S13301 collagen alpha 1(V |
| 14 | 160.5 | 18.0 | 744 | 2 | S15435 collagen alpha 1(X |
| 15 | 155.5 | 17.4 | 246 | 2 | S29328 complement subcomp |
| 16 | 154.5 | 17.3 | 743 | 1 | S23779 collagen alpha 1(V |
| 17 | 153 | 17.1 | 193 | 2 | A37873 cerebellin precurs |
| 18 | 150.5 | 16.8 | 253 | 2 | I49560 complement C1q B c |
| 19 | 144 | 16.1 | 215 | 2 | C48150 hibernation-relate |
| 20 | 143 | 16.0 | 196 | 2 | A48150 hibernation-relate |
| 21 | 135.5 | 15.2 | 245 | 1 | CIHUOA complement subcomp |
| 22 | 134 | 15.0 | 245 | 2 | S19018 complement subcomp |
| 23 | 132.5 | 14.8 | 635 | 2 | A57131 collagen alpha 2(V |
| 24 | 129.5 | 14.5 | 170 | 2 | B57131 collagen alpha 2(V |
| 25 | 129 | 14.4 | 992 | 2 | T08772 hypothetical prote |
| 26 | 123.5 | 13.8 | 423 | 2 | A55797 collagen precursor |
| 27 | 85 | 9.5 | 589 | 2 | H83085 conserved hypothet |
| 28 | 85 | 9.5 | 1228 | 2 | A57384 multimerin, endoth |
| 29 | 82.5 | 9.2 | 472 | 2 | E98020 hypothetical prote |

| | | | | | | |
|----|------|-----|------|---|--------|--------------------|
| 30 | 82.5 | 9.2 | 1433 | 2 | A46053 | bullous pemphigoid |
| 31 | 81.5 | 9.1 | 420 | 2 | E75130 | hypothetical prote |
| 32 | 81 | 9.1 | 618 | 2 | T24228 | hypothetical prote |
| 33 | 80.5 | 9.0 | 216 | 2 | T29039 | hypothetical prote |
| 34 | 80 | 8.9 | 657 | 2 | E70025 | probable acylamino |
| 35 | 78.5 | 8.8 | 602 | 2 | T35782 | probable secreted |
| 36 | 76.5 | 8.6 | 425 | 2 | T41172 | hypothetical prote |
| 37 | 76.5 | 8.6 | 706 | 2 | S33761 | transferrin precur |
| 38 | 76 | 8.5 | 2796 | 2 | JC4743 | fatty-acid synthas |
| 39 | 75.5 | 8.4 | 175 | 2 | B83411 | hypothetical prote |
| 40 | 75.5 | 8.4 | 250 | 2 | S27918 | coat protein - gar |
| 41 | 75.5 | 8.4 | 250 | 2 | JQ2181 | viral coat 27.4K p |
| 42 | 75.5 | 8.4 | 400 | 2 | B72650 | hypothetical prote |
| 43 | 75.5 | 8.4 | 493 | 2 | E71725 | glutamyl-tRNA amid |
| 44 | 75.5 | 8.4 | 660 | 2 | T03038 | probable inhibitor |
| 45 | 75.5 | 8.4 | 1445 | 2 | A59437 | KIAA1204 protein { |

ALIGNMENTS

RESULT 1
T14782
hypothetical protein DKFZp586B0621.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Sep-2000
C;Accession: T14782
R;Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, August 1999
A;Reference number: Z18184
A;Accession: T14782
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-219 <OTT>
A;Cross-references: EMBL:AL110261
A;Experimental source: adult uterus; clone DKFZp586B0621
C;Genetics:
A;Note: DKFZp586B0621.1
C;Superfamily: complement C1q carboxyl-terminal homology

| | | | | |
|-----------------------|--------|----------------------------------|---|-------------------------------|
| Query Match | 23.3% | Score 208.5; | DB 2; | Length 219; |
| Best Local Similarity | 35.7%; | Pred. No. 1.4e-12; | | |
| Matches | 56; | Conservative 18; | Mismatches 72; | Indels 11; Gaps 3; |
| QY | 4 | DAPARGPPAP-----EPRSASFSAARTSLVGS | DAGPGRHQPLAFPTFEVNI | GDFDAA 58 |
| Db | 61 | EAGPAGPTGPAGECSP | PPRSASFSAKRSSESRV-----PPPSDAPLPFDRVLVNEQHYDAV | 115 |
| QY | 59 | AGVFRCLPGAYPF | FTLGKLPKRTLSVKLMKNRDE | VQAMITYDDGASRRREMOSQSVMLA 118 |
| Db | 116 | TGKFTCQVP | GVYFFA-VHATVYRASLQFDLVKNGESIASFFQFGGNPRPASLSGAMVR | 174 |
| QY | 119 | LRRGDVWL | LSHHDG | YGAYSNHGKITYFGFLVYPD 155 |
| Db | 175 | LEPEDQVWVQVGV | GDYIGIYASIKTDS | TFSGFLVYSD 211 |

RESULT 2
CIHUOB
complement subcomponent C1q chain B precursor [validated] - human
N;Alternate names: complement subcomponent C1q beta chain
C;Species: Homo sapiens (man)
C;Date: 22-May-1981 #sequence_revision 31-May-1996 #text_change 08-Dec-2000
C;Accession: B23422; A23422; B90304; A90301; B90315; A03206
R;Reid, K.B.M.
Biochem. J. 231, 729-735, 1985
A;Title: Molecular cloning and characterization of the complementary DNA and gene coding
A;Reference number: A23422; MUID:86076906; PMID:3000358
A;Accession: B23422
A;Molecule type: DNA
A;Residues: 'HS', 1-32 <REI>
A;Note: the authors translated the codon AGR for the second position as Arg; they were
A;Accession: A23422

A/Molecule type: mRNA
A/Residues: 28-253 <RE1>
A/Cross-references: EMBL:X03084
A/Note: the authors translated the codon ACA for residue 46 as Ile
R/Reid, K.B.M.
Biochem. J. 179, 367-371, 1979
A/Title: Complete amino acid sequences of the three collagen-like regions present in sub
A/Reference number: A90304; MUID:80020137; PMID:486087
A/Accession: B90304
A/Molecule type: protein
A/Residues: 'E', 29-84, 'D', 86-99, 'P', 101-135 <RE5>
R/Reid, K.B.M.; Thompson, E.O.P.
Biochem. J. 173, 863-868, 1978
A/Title: Amino acid sequence of the N-terminal 108 amino acid residues of the B chain of
A/Reference number: A90301; MUID:79041552; PMID:708376
A/Accession: A90301
A/Molecule type: protein
A/Residues: 28-99, 'P', 101-195 <RE3>
R/Reid, K.B.M.; Gagnon, J.; Frampton, J.
Biochem. J. 203, 559-569, 1982
A/Title: Completion of the amino acid sequences of the A and B chains of subcomponent C1
A/Reference number: A90315; MUID:82283890; PMID:6981411
A/Accession: B90315
A/Molecule type: protein
A/Residues: 136-253 <RE4>
A/Note: 176-Glx may also be present
C/Comment: The first component of complement is a calcium-dependent complex of the three
ivation of C1r (enzyme), C1s (proenzyme), and the other eight components of complement.
C/Comment: The C1q subcomponent is composed of nine subunits, six of which are disulfide
(see PIR:C1HUQC) chain. Equimolar amounts of the A, B, and C chains are found after red
C/Genetics:
A/Gene: GDB:C1QB
A/Cross-references: GDB:119043; OMIM:120570
A/Map position: 1p36.3-1p34.1
C/Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom
C/Keywords: complement pathway; glycoprotein; heterodimer; hydroxyllysine; hydroxyproline
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-253/Product: complement subcomponent C1q chain B #status experimental <MAT>
F:33-116/Domain: collagenous, triple helix <COL>
F:123-249/Domain: complement C1q carboxyl-terminal homology <C1Q>
F:28/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F:31/Disulfide bonds: interchain (to chain A-26) #status experimental
F:35, 38, 41, 53, 56, 65, 83, 86, 101, 104, 107/Modified site: 4-hydroxyproline (Pro) #status expe
F:59, 62, 77, 92, 98, 110/Modified site: 5-hydroxyllysine (Lys) #status experimental
F:59, 62, 98, 110/Binding site: carbohydrate (Lys) (covalent) #status experimental
Query Match 20.4%; Score 182; DB 1; Length 253;
Best Local Similarity 29.1%; Pred. NO. 5.8e-10;
Matches 48; Conservative 32; Mismatches 61; Indels 24; Gaps 5;
QY 6 PARPPAPPP-----RSAFSAARTSLVSGDAGPPRHQPLAFDTEFVNIG 52
DB 97 PKGGPAGPAGPAGPKGESGDYKATQKIAFSATRTINV-----PLRRDQITIRFDHVTMMN 150
QY 53 GDFDAAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMT-YDDGASRRREMÖ 111
DB 151 NNYPEPRSGKFTCKVPGLYFTYHASS--RGNLCVNLMRGRERAQKVYFCDYAYNTFQVT 208
QY 112 SQSVMLALRRGDAVWLLSHDHGYGAYSNHGXYTFSGFLVYPDL 156
DB 209 TGGMVLKLEGGENVFLQATDKN--SLIGMEGANSIFSGFLFPDM 251
RESULT 3
B48150
hibernation-related protein HP-25 precursor - Siberian chipmunk
C/Species: Eutamias sibiricus (Siberian chipmunk)
C/Date: 16-Feb-1994 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C/Accession: B48150; B41752
R/Takamatsu, N.; Ohba, K.; Kondo, J.; Kondo, N.; Shiba, T.
Mol. Cell. Biol. 13, 1516-1521, 1993
A/Title: Hibernation-associated gene regulation of plasma proteins with a collagen-like
A/Reference number: A48150; MUID:93180798; PMID:8441393

A/Accession: B48150
A/Molecule type: mRNA; protein
A/Residues: 1-215 <TAK>
A/Cross-references: GB:D12975; NID:9287469; PIDN:BAA02352.1; PID:9287470
A/Note: the source is designated as Tamias asiaticus in Genbank entry TMSHP25, release
A/Note: sequence extracted from NCBI backbone (NCBIN:125946, NCBI:P:125947)
R/Kondo, N.; Kondo, J.
J. Biol. Chem. 267, 473-478, 1992
A/Title: Identification of novel blood proteins specific for mammalian hibernation.
A/Reference number: A41752; MUID:92112696; PMID:1730610
A/Accession: B41752
A/Status: preliminary
A/Molecule type: protein
A/Residues: 29-62, 84-130, 172-183, 187-192, 201-215 <KON>
C/Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom
C/Keywords: glycoprotein; hibernation; plasma
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-215/Product: hibernation-related protein HP-25 #status experimental <MAT>
F:91-214/Domain: complement C1q carboxyl-terminal homology <C1Q>
F:167/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 19.9%; Score 177.5; DB 2; Length 215;
Best Local Similarity 32.4%; Pred. NO. 1.3e-09;
Matches 58; Conservative 21; Mismatches 49; Indels 51; Gaps 10;
QY 5 AP-ARGPPAP-----EPBSAFSAARTSLVSGDAGPPRHQPLA 43
DB 56 APGALGPPGPPGVPGIPRPGPPGDVEKCSSRPSAFVAVKL-----SERPPF-FQPIV 108
QY 44 FDTFVNIGDGFDAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMTYDDG 103
DB 109 FKEALYNQEGHFNMATGEFSCVLPGVYNTFGFDI-RLFQSSVKIRLM--RDGIQV----- 159
QY 104 ASRRREMÖSQ-----SVMLALRRGDAVWLLSHDHGYGAYSNHG-KYTFSGFLVY 153
DB 160 --REKEAQANDSYKHAMGSVIMALGKDKWLESKLR--GTESEKGIITHIVFGYLLY 213
RESULT 4
JC4708
gelatin-binding 28K protein precursor - human
N/Alternate names: adipose specific collagen-like factor
C/Species: Homo sapiens (man)
C/Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 20-Sep-1999
C/Accession: JC4708; JC4944
R/Maeda, K.; Okubo, K.; Shimomura, I.; Funahashi, T.; Matsuzawa, Y.; Matsubara, K.
Biochem. Biophys. Res. Commun. 221, 286-289, 1996
A/Title: cDNA cloning and expression of a novel adipose specific collagen-like factor,
A/Reference number: JC4708; MUID:962241171; PMID:8619847
A/Accession: JC4708
A/Molecule type: mRNA
A/Residues: 1-244 <MAE>
A/Cross-references: DDBJ:D45371; NID:9871886; PIDN:BAA08227.1; PID:9871887
A/Experimental source: adipose tissue
R/Nakano, Y.; Tobe, T.; Choi-Miura, N.H.; Mazda, T.; Tomita, M.
J. Biochem. 120, 803-812, 1996
A/Title: Isolation and characterization of GBP28, a novel gelatin-binding protein purif
A/Reference number: JC4944; MUID:97103474; PMID:8947845
A/Accession: JC4944
A/Molecule type: protein
A/Residues: 19-38, 93-100, 101-112, 135-149, 173-178 <NAK>
C/Comment: This protein is an endogenous factor that binds with a collagen-like domain
C/Genetics:
A/Gene: apM1
C/Superfamily: unassigned collagens; complement C1q carboxyl-terminal homology
C/Keywords: adipose tissue; glycoprotein; hydroxyproline
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-244/Product: gelatin-binding 28kDa protein #status experimental <MAT>
F:42-107/Region: collagen-like
F:114-241/Domain: complement C1q carboxyl-terminal homology <C1Q>
F:95/Modified site: 4-hydroxyproline (Pro) #status experimental
F:230/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.6%; Score 175; DB 2; Length 244;
Best Local Similarity 31.1%; Pred. No. 2.6e-09;
Matches 51; Conservative 24; Mismatches 65; Indels 24; Gaps 6;

QY 7 ARGPAPP-----EP-----RSAFSAARTSLVGSADGPRHQPAPDTEFVNIG 52
DB 88 AEGPRGPGIQRKGEPEGAYVRSAPS-----VGLTYVTIPNMPIRFKIFYNQ 140
QY 53 GDFDAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMTYDDGASRRREMOS 112
DB 141 NHYDSTGKFHCNIPGLYFYAYHI-TVMKDVKVSIFK-KDKAMLFTYDQYQENNVQAS 198
QY 113 QSVMLALRRGDAVWLISH-DHDGYGAYSNHGKYTFSGFLVYPD 155
DB 199 GSVLHLLEVGDQVWLQVYGEGERNGLYADNDNDSTFTGLLYHD 242

RESULT 5
C1HUQC
complement subcomponent C1q chain C precursor - human
N/Alternate names: complement subcomponent C1q gamma chain
C/Species: Homo sapiens (man)
C/Date: 22-May-1981 #sequence_revision 31-May-1996 #text_change 22-May-1998
C/Accession: S14351; A03207
R/Seller, G.C.; Blake, D.J.; Reid, K.B.M.
Biochem. J. 274, 481-490, 1991
A/Title: Characterization and organization of the genes encoding the A-, B- and C-chains
A/Reference number: S14350; MUID:91174759; PMID:1706597
A/Accession: S14351
A/Status: not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-245 <SEL>
R/Reid, K.B.M.
Biochem. J. 179, 367-371, 1979
A/Title: Complete amino acid sequences of the three collagen-like regions present in sub
A/Reference number: A90304; MUID:80020137; PMID:486087
A/Accession: A03207
A/Molecule type: protein
A/Residues: 29-56, 'P', 58-65, 'K', 67-71, 'P', 73-83, 'K', 85-86, 'D', 88-89, 'N', 91-122 <REI>
C/Comment: The first component of complement is a calcium-dependent complex of the three
ivation of C1r (enzyme), C1s (proenzyme), and the other eight components of complement.
C/Comment: The C1q subcomponent is composed of nine subunits, six of which are disulfide
dimers of the C chain. Equimolar amounts of the A, B, and C chains are found after redu
C/Genetics:
A/Gene: GDB:C1QG
A/Cross-references: GDB:128132; OMIM:120575
A/Map position: 1p36.3-1p34.1
A/Introns: 60/3
C/Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom
C/Keywords: complement pathway; glycoprotein; homodimer; hydroxyllysine; hydroxyproline;
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-245/Product: complement subcomponent C1q chain B #status predicted <MAT>
F:31-114/Domain: collagenous, triple helix <COL>
F:121-244/Domain: complement C1q carboxyl-terminal homology <C1Q>
F:32/Disulfide bonds: interchain #status experimental
F:36,39,42,45,54,63,81,93,96,99,105/Modified site: 4-hydroxyproline (Pro) #status experi
F:57,72,75/Modified site: 5-hydroxyllysine (Lys) #status experimental
F:75/Binding site: carbohydrate (Lys) (covalent) #status experimental

Query Match 19.0%; Score 170; DB 1; Length 245;
Best Local Similarity 26.6%; Pred. No. 7.9e-09;
Matches 46; Conservative 27; Mismatches 56; Indels 44; Gaps 6;

QY 9 GPPA-----PPEP-----RSAFSAARTSLVGSADGPRHQP-----L 42
DB 91 GPPGMPGVPGMGIPEGEBEGRYKQKFGSVFTVTR-----QTHQPPAPNSLI 138
QY 43 AFDTEFVNIGDFDAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMTYDD 102
DB 139 RFNAVLTPQGDYDTSTGKFTCKVPGLYFYHAS---HTANLCVLLYRSGVKVTFPCG 194
QY 103 GASRRREMOSQSVMLALRRGDAVWLISHDHDGYGAYSNHGKYTFSGFLVYPD 155

DB 195 HTSKTNQVNSGGVLLRLQVGEVWLAVNDY--YDMVGIOGSDSVFSGFLFPD 245

RESULT 6
S23297
collagen alpha 1(X) chain precursor - chicken
N/Alternate names: type X collagen
C/Species: Gallus gallus (chicken)
C/Date: 07-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 13-Aug-1999
C/Accession: S23297; A31896; S65594; S77711; I50218
R/Ninomiya, Y.; Castagnola, P.; Gerecke, D.; Gordon, M.K.; Jacenko, O.; LuValle, P.; Mc
maguchi, N.; Olsen, B.R.
in Extracellular Matrix Genes, Sandell L.J. and Boyd C.D., eds., pp.79-114, Academic Pr
A/Title: The molecular biology of collagens with short triple-helical domains.
A/Reference number: S22243
A/Accession: S23297
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-674 <NIN>
R/LuValle, P.; Ninomiya, Y.; Rosenblum, N.D.; Olsen, B.R.
J. Biol. Chem. 263, 18378-18385, 1988
A/Title: The type X collagen gene. Intron sequences split the 5'-untranslated region an
A/Reference number: A31896; MUID:89054019; PMID:2461368
A/Accession: A31896
A/Molecule type: mRNA
A/Residues: 1-75 <LUV>
R/Ninomiya, Y.; Gordon, M.; van der Rest, M.; Schmid, T.; Linsenmayer, T.; Olsen, B.R.
J. Biol. Chem. 261, 5041-5050, 1986
A/Title: The developmentally regulated type X collagen gene contains a long open readin
A/Reference number: I50218; MUID:86168227; PMID:3082876
A/Accession: S65594
A/Molecule type: DNA
A/Residues: 'T', 9, 'D', 11-12, 'EDQMKLYLFTM', 30-31, 'TCKSGRAFTTYMILQNMADLVSSHT', 48-89, 'L',
629, 'PQAVLSLISWRITKCSSCQIQNPMVSIPLNMFILLSQVSYLKSNNIPTMS' <NIN1>
A/Cross-references: EMBL:M13496; NID:g211699; PIDN:AAA48736.1; PID:g211700
A/Accession: S77711
A/Molecule type: protein
A/Residues: 104-112, 'X', 114-117; 453-466 <NIN2>
C/Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
C/Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer; hydroxyproline,
F:1-18/Domain: signal sequence #status predicted <SIG>
F:547-673/Domain: complement C1q carboxyl-terminal homology <C1Q>
F:453,456/Modified site: hydroxyproline (Pro) #status experimental
F:611/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 18.6%; Score 166.5; DB 2; Length 674;
Best Local Similarity 28.9%; Pred. No. 5.8e-08;
Matches 50; Conservative 19; Mismatches 71; Indels 33; Gaps 5;

QY 9 GPPAPPEP-----RSAFSAARTSLVGSADGPRHQP-----SDAGPGPR 38
DB 501 GPPGPPPGQSTIPEGYVKGESRELSGMSFMKAGANQALTGMPVSAFTVILSKAYPGAT 560
QY 39 HQPLAFDTEFVNIGDFDAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAM 98
DB 561 -VPKPKDKILYNRQGHYDPRFTGICRIIPGLYFYSYHV-HAKGTNWWVALYKNGSPVM-Y 617
QY 99 IYDDGASRRREMOSQSVMLALRRGDAVWLISHDHDGYGAYSNHGKYTFSGFL 151
DB 618 TYDEYOKGYLDQASGSAVIDLMENDQVWLQLPNSESNGLYSSEYVHSSFSFGFL 670

RESULT 7
S49158
complement protein C1q beta chain precursor - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 20-Aug-1999
C/Accession: S49158
R/Schwaible, W.; Petry, F.; Loos, M.
submitted to the EMBL Data Library, March 1993
A/Description: cDNA cloning and expression of the mRNA encoding the B-chain of rat C1q.
A/Reference number: S49158
A/Accession: S49158

Query Match 18.3%; Score 164; DB 1; Length 680;
Best Local Similarity 25.6%; Pred. No. 1e-07;
Matches 50; Conservative 20; Mismatches 71; Indels 54; Gaps 4;

QY 9 GPPAPPEPRSAFSAARTSLVSGDAGPGPRHQ----- 40
Db 490 GPPGPPGPR---GHSGEPLPGPPGPPGQAVMPEGFIKAGQRPRLSGMPLVSNHGV 546
QY 41 -----PLAFDTEFVNIGGDFDAAGVFRCLPGAYFFSFTLGKL 79
Db 547 TGMPSAFTVILSKAYPAIGTFPFEDKILYNROQHYPRTGIFTCCIPGIYFSYHV-HV 605
QY 80 PRKTLVSKLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRGDAVWLSSHHDGCGAYS 139
Db 606 KGTWVWGLYKNGTPVM-YTYDEYTKGYLDQASGSAIDLTEHQVWLQLPNAESNGLYS 664
QY 140 NHGKYITFSGFLVYP 154
Db 665 SEYVHSSFSGLVAP 679

RESULT 10
S31216
collagen alpha 1(X) chain precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text change 13-Aug-1999
C/Accession: S31216; S28807; S22215; S30127; I48299; S26397; S31830
R/Kong, R.Y.C.; Kwan, K.M.; Lau, E.T.; Thomas, J.T.; Boot-Handford, R.P.; Grant, M.E.; C
Eur. J. Biochem. 213, 99-111, 1993
A/Title: Intron-exon structure, alternative use of promoter and expression of the mouse
A/Reference number: S31216; MUID:93238750; PMID:8477738
A/Accession: S31216
A/Molecule type: DNA
A/Residues: 1-680 <KON>
A/Cross-references: EMBL:Z21610; NID:g49793; PIDN:CAA79736.1; PID:g49794
R/Elima, K.; Eerola, I.; Rosati, R.; Metsaeranta, M.; Garofalo, S.; Peraelae, M.; de Cro
Biochem. J. 289, 247-253, 1993
A/Title: The mouse collagen X gene: complete nucleotide sequence, exon structure and exp
A/Reference number: S28807; MUID:93143676; PMID:8424763
A/Accession: S28807
A/Molecule type: DNA
A/Residues: 1-285, 'A', 287-680 <ELI>
A/Cross-references: EMBL:X67348; NID:g50480; PIDN:CAA47763.1; PID:g50481
R/Elima, K.; Metsaeranta, M.; Kallio, J.; Peraelae, M.; Eerola, I.; Garofalo, S.; de Cro
Biochim. Biophys. Acta 1130, 78-80, 1992
A/Title: Specific hybridization probes for mouse alpha-2(IX) and alpha-1(X) collagen mRN
A/Reference number: S22215; MUID:92182017; PMID:1543751
A/Accession: S22215
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 385-450, 'K', 452-627 <ELA>
A/Cross-references: EMBL:X63013; NID:g49795; PIDN:CAA44741.1; PID:g49796
R/Apte, S.S.; Olsen, B.R.
Matrix 13, 165-179, 1993
A/Title: Characterization of the mouse type X collagen gene.
A/Reference number: S30127; MUID:93261348; PMID:8492743
A/Accession: S30127
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-12, 'F', 14-26, 'S', 28-247, 'L', 249-285, 'A', 287-305, 'F', 307-416, 'S', 418-499, 'L
Eur. J. Biochem. 206, 217-224, 1992
A/Title: Cloning of the human and mouse type X collagen genes and mapping of the mouse t
A/Reference number: I48299; MUID:92267014; PMID:1587271
A/Accession: I48299
A/Status: preliminary; translated from GB/EMBL/DDbJ
A/Molecule type: DNA
A/Residues: 52-247, 'L', 249-285, 'A', 287-305, 'F', 307-416, 'S', 418-499, 'L', 501-566, 'C', 568, '
A/Cross-references: EMBL:X65121; NID:g50482; PIDN:CAA46237.1; PID:g667031
R/Summers, T.A.; Irwin, M.H.; Mayne, R.; Balian, G.
J. Biol. Chem. 263, 581-587, 1988
A/Title: Monoclonal antibodies to type X collagen. Biosynthetic studies using an antibod
A/Reference number: S26397; MUID:88087150; PMID:2826450

A/Accession: S26397
A/Molecule type: protein
A/Residues: 'SDGYFSQ', 24-26, 'KQ' <SUM>
C/Genetics:
A/Gene: Col10a-1
A/Map position: 10
A/Introns: 51/3
C/Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
C/Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-680/Product: collagen alpha 1(X) chain #status predicted <MAT>
F:553-679/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 18.3%; Score 164; DB 2; Length 680;
Best Local Similarity 25.1%; Pred. No. 1e-07;
Matches 49; Conservative 22; Mismatches 70; Indels 54; Gaps 4;

QY 9 GPPAPPEPRSAFSAARTSLVSGDAGPGPRHQ----- 40
Db 490 GPPGPPGPR---GHSGEPLPGPPGPPGQAVMPEGFIKAGQRPRLSGMPLVSNHGV 546
QY 41 -----PLAFDTEFVNIGGDFDAAGVFRCLPGAYFFSFTLGKL 79
Db 547 TGMPSAFTVILSKAYPAVGAPIPFDEILYNROQHYPDRSGIFTCKIPGIYFSYHV-HV 605
QY 80 PRKTLVSKLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRGDAVWLSSHHDGCGAYS 139
Db 606 KGTWVWGLYKNGTPM-YTYDEYSKGYLDQASGSAIMELTEHQVWLQLPNAESNGLYS 664
QY 140 NHGKYITFSGFLVYP 154
Db 665 SEYVHSSFSGLVAP 679

RESULT 11
A34246
collagen alpha 1(VIII) chain precursor - rabbit
C/Species: Oryctolagus cuniculus (domestic rabbit)
C/Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 10-Sep-1999
C/Accession: A34246
R/Yamaguchi, N.; Benya, P.D.; van der Rest, M.; Ninomiya, Y.
J. Biol. Chem. 264, 16022-16029, 1989
A/Title: The cloning and sequencing of alpha1(VIII) collagen cDNAs demonstrate that typ
omains similar to those of type X collagen.
A/Reference number: A34246; MUID:89380199; PMID:2476437
A/Accession: A34246
A/Molecule type: mRNA
A/Residues: 1-744 <YAM>
A/Cross-references: GB:J05042; NID:g164895; PIDN:AAA31204.1; PID:g164896
C/Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-744/Product: collagen alpha 1(VIII) chain #status predicted <MAT>
F:21-117/Region: amino-terminal nonhelical
F:118-571/Region: interrupted helical
F:572-744/Region: carboxyl-terminal nonhelical
F:617-743/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 18.3%; Score 163.5; DB 1; Length 744;
Best Local Similarity 25.0%; Pred. No. 1.3e-07;
Matches 47; Conservative 26; Mismatches 66; Indels 49; Gaps 5;

QY 9 GPPAPPEPRSAFSAARTSLVSGDAGPG----- 36
Db 563 GPPGPPGPPAVMPPTPAQGEYLPDMGLCIDGVKTPHAYAAKK-----GKNGPAYEMPA 617
QY 37 -----PRHQPLAFDTEFVNIGGDFDAAGVFRCLPGAYFFSFTLGKLPKRLTSLV 86
Db 618 FTAELTAPPPVGAPIKFDRLLYNGRQYNPQTGIFTCEVPGVYFAYHV-HCKGQNVWV 676
QY 87 KLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRGDAVWLSSHHDGCGAYSNHGKYIT 146
Db 677 ALFKNNEPVM-YTYDEYKKGFLDQASGSAVLLLRPGDRVFLQMPSEQAAGLYAGQYVHSS 735

A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-246 <PET>
A;Cross-references: EMBL:X66295; NID:g50228; PIDN:CAA46993.1; PID:g50229
C;Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom
F;122-245/Domain: complement C1q carboxyl-terminal homology <C1Q>

```

Query Match      17.4%; Score 155.5; DB 2; Length 246;
Best Local Similarity 24.4%; Pred. No. 1.9e-07;
Matches 44; Conservative 26; Mismatches 69; Indels 41; Gaps 4;

QY      8 RGPAPPEPRSAFSAARTSLVSGDAGPPRHQP----- 41
      ||| | | | | | | | | | |
Db      76 KGEPMGHRGKNGPRGTGLPGD---PGRGPPGEPGVEGRYKQKHQSVFTVTRQTTQY 132
      ||| | | | | | | | | | |

QY      42 -----LAFDTEFVNIGGDFDAAGVFRCLPGAYFFSFTLGLKLPKRLSVKLMKNRDEV 95
      :|:|:| | | | | | | | | | |

Db      133 PEANALVRFNSVVTNPQGHYNPSTGKFTCEVPGLYFYV---YTSHTANLCVHLNLTLA 188
      :|:|:| | | | | | | | | | |

QY      96 QAMITYDGDASRRREMOSQSVMLLRGDVAVLLSHDHGDYGAYSNHGKYITFSGFLVYPD 155
      :|:|:| | | | | | | | | | |

Db      189 RVASFCDHMFNSKQVSSGALLRLQRGDDEVWLSVNDYN--GMVGIEGSNSVFSGFLLPD 246
      :|:|:| | | | | | | | | | |

```

Search completed: January 12, 2004, 08:16:46
Job time : 11.8403 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:09 ; Search time 6.10778 Seconds
(without alignments)
1308.910 Million cell updates/sec

Title: US-10-085-167-2_COPY_160_358
Perfect score: 894
Sequence: 1 ADADAPARGPPAPPEPRSAF.....LVYPDLAPAPGLGASELL 170

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|---------------------|
| 1 | 894 | 100.0 | 329 | 1 CQT4_HUMAN | Q9bxj3 homo sapien |
| 2 | 208.5 | 23.3 | 243 | 1 CQT5_HUMAN | Q9bxj0 homo sapien |
| 3 | 202.5 | 22.7 | 281 | 1 CQT1_HUMAN | Q9bxj1 homo sapien |
| 4 | 194 | 21.7 | 278 | 1 CQT6_HUMAN | Q9bxj9 homo sapien |
| 5 | 192.5 | 21.5 | 258 | 1 C1RF_MOUSE | Q88992 mus musculu |
| 6 | 189.5 | 21.2 | 258 | 1 C1RF_HUMAN | Q75973 homo sapien |
| 7 | 188 | 21.0 | 247 | 1 APM1_MOUSE | Q60994 mus musculu |
| 8 | 187.5 | 21.0 | 285 | 1 CQT2_HUMAN | Q9bxj5 homo sapien |
| 9 | 185.5 | 20.7 | 255 | 1 GLIC_MOUSE | Q9esn4 mus musculu |
| 10 | 182 | 20.4 | 251 | 1 C1QB_HUMAN | P02746 homo sapien |
| 11 | 177.5 | 19.9 | 215 | 1 HP25_TAMSI | Q06576 tamias sibi |
| 12 | 175 | 19.6 | 244 | 1 APM1_HUMAN | Q15848 homo sapien |
| 13 | 171 | 19.1 | 419 | 1 COLE_LEPMA | P98085 leopomis mac |
| 14 | 170 | 19.0 | 245 | 1 C1QC_MOUSE | P02747 homo sapien |
| 15 | 166.5 | 18.6 | 674 | 1 CA1A_CHICK | P08125 gallus gall |
| 16 | 165.5 | 18.5 | 253 | 1 C1QB_RAT | P31721 rattus norv |
| 17 | 164 | 18.3 | 680 | 1 CA1A_HUMAN | Q03692 homo sapien |
| 18 | 164 | 18.3 | 680 | 1 CA1A_MOUSE | Q05306 mus musculu |
| 19 | 163.5 | 18.3 | 744 | 1 CA18_RABIT | P14282 oryctolagus |
| 20 | 163 | 18.2 | 224 | 1 CERL_RAT | P98087 rattus norv |
| 21 | 163 | 18.2 | 674 | 1 CA1A_BOVIN | P23206 bos taurus |
| 22 | 162 | 18.1 | 744 | 1 CA18_HUMAN | P27658 homo sapien |
| 23 | 159 | 17.8 | 246 | 1 CQT3_HUMAN | Q9bxj4 homo sapien |
| 24 | 155.5 | 17.4 | 246 | 1 C1QC_MOUSE | Q02105 mus musculu |
| 25 | 155.5 | 17.4 | 743 | 1 CA18_MOUSE | Q00780 mus musculu |
| 26 | 153 | 17.1 | 193 | 1 CERB_HUMAN | P23435 homo sapien |
| 27 | 153 | 17.1 | 193 | 1 CERB_MOUSE | Q9r171 mus musculu |
| 28 | 150.5 | 16.8 | 253 | 1 C1QB_MOUSE | P14106 mus musculu |
| 29 | 147.5 | 16.5 | 508 | 1 OTOL_ONCKE | P83371 oncorhynch |
| 30 | 147 | 16.4 | 289 | 1 CQT7_HUMAN | Q9bxj2 homo sapien |
| 31 | 144 | 16.1 | 215 | 1 HP27_TAMSI | Q06577 tamias sibi |
| 32 | 143 | 16.0 | 196 | 1 HP20_TAMSI | Q06575 tamias sibi |
| 33 | 135.5 | 15.2 | 245 | 1 C1QA_HUMAN | P02745 homo sapien |

| | | | | | |
|----|-------|------|------|--------------|--------------------|
| 34 | 134 | 15.0 | 245 | 1 C1QA_MOUSE | P98086 mus musculu |
| 35 | 132.5 | 14.8 | 635 | 1 CA28_HUMAN | P25067 homo sapien |
| 36 | 129.5 | 14.5 | 170 | 1 CA28_MOUSE | P25318 mus musculu |
| 37 | 126.5 | 14.1 | 201 | 1 CERL_HUMAN | Q9ntu7 homo sapien |
| 38 | 89.5 | 10.0 | 265 | 1 Y176_HUMAN | Q14681 homo sapien |
| 39 | 89 | 10.0 | 384 | 1 DUS9_HUMAN | Q99956 homo sapien |
| 40 | 85 | 9.5 | 1228 | 1 ECM_HUMAN | Q13201 homo sapien |
| 41 | 80 | 8.9 | 657 | 1 YUXL_BACSU | P39839 bacillus su |
| 42 | 77.5 | 8.7 | 664 | 1 KCCB_HUMAN | Q13554 homo sapien |
| 43 | 76.5 | 8.6 | 706 | 1 TRFE_HORSE | P27425 equus cabal |
| 44 | 76.5 | 8.6 | 1286 | 1 SMC4_MOUSE | Q8CG47 mus musculu |
| 45 | 75.5 | 8.4 | 493 | 1 GATA_RICPR | Q9ze10 rickettsia |

ALIGNMENTS

RESULT 1

CC CQT4_HUMAN STANDARD; PRT; 329 AA.
AC Q9BXJ3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Complement-clq tumor necrosis factor-related protein 4 precursor.
GN C1QTNF4 OR CTRP4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Holloway J.L., Lok S.;
RT "Homo sapiens complement-clq tumor necrosis factor-related protein.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 2 C1Q domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF329838; AAK17962.1; -
DR Genew; HGNC:14346; C1QTNF4.
DR InterPro; IPR001073; C1q.
DR Pfam; PF00386; C1q; 2.
DR SMART; SM00110; C1Q; 2.
DR PROSITE; PS01113; C1Q; 2.
KW Repeat; Signal.

FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 329 COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT DOMAIN 23 159 RELATED PROTEIN 4.
FT FT 170 C1Q 1.
SQ SEQUENCE 329 AA; 35265 MW; 331C7DBF26036915 CRC64;

Query Match 100.0%; Score 894; DB 1; Length 329;
Best Local Similarity 100.0%; Pred. No. 1.1e-76;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| QY | 1 | ADADAPARGPPAPPEPRSAARTRSLVSGDAGPGPRHQLAFDTEFVNIGGDFDAAG | 60 |
| DB | 160 | ADADAPARGPPAPPEPRSAARTRSLVSGDAGPGPRHQLAFDTEFVNIGGDFDAAG | 219 |
| QY | 61 | VFRGRLPGAYFFSFTLGKLPKRLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALR | 120 |
| DB | 220 | VFRGRLPGAYFFSFTLGKLPKRLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALR | 279 |
| QY | 121 | RGDAVWLSHDHDGYGAYSNHGKYITFSGLVYPDLAPAPGLGASELL | 170 |

Db 280 RGDVWLLSHDHGCGAYSNHGKTYTFSGFLVYPDLAPAPGLGASELL 329

```
RESULT 2
COT5 HUMAN STANDARD; PRT; 243 AA.
ID COT5 HUMAN
AC Q9BXJ0; Q9UF4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Complement-c1q tumor necrosis factor-related protein 5 precursor.
GN C1QTNF5 OR CTRP5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Sheppard P.O., Humes J.M.;
RT "Homo sapiens complement-c1q tumor necrosis factor-related protein.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 25-243 FROM N.A.
RC TISSUE=Uterus;
RA Ottenwaelder B., Obermaier B., Mewes H.-W., Gassenhuber J.,
RA Wiemann S.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C1Q domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AF329841; AA017965.1;
DR EMBL; AL110261; CAB53702.1;
DR PIR; T14782; T14782.
DR Genew; HGNC:14344; C1QTNF5.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; FALSE_NEG.
KW Collagen; Signal.
FT SIGNAL 1 15
FT CHAIN 16 243
FT FT 16 243
FT FT 30 95
FT DOMAIN 97 243
SQ SEQUENCE 243 AA; 25298 MW; 7CCDA65CDA7EB784 CRC64;

Query Match 23.3%; Score 208.5; DB 1; Length 243;
Best Local Similarity 35.7%; Pred. No. 1.7e-12;
Matches 56; Conservative 18; Mismatches 72; Indels 11; Gaps 3;

QY 4 DAPARGPPAP-----EPRSAFSAARTSLVGSADAGPGRHQPLAFDTFVNIGGDFDAA 58
DB 85 EAGPAGPTGPAGECSVPPRSAFSAKRSERV-----PPSDAPLPFDRVLVNEQGHYDAV 139
QY 59 AGVFRCLPGAYFFSFTLGLPKRLSVKLMKNRDEVQAMTYDDGASRRRQMSQSVMLA 118
DB 140 TGFCTCQVPGVYFA-VHATVYRASLQFDLVKNGESIASFFQFFGGMKPKASLSGAMVR 198
QY 119 LRRGDVWLLSHDHGCGAYSNHGKTYTFSGFLVYPD 155
DB 199 LEPEDQVWVQVGVDYIGIYASIKTDSFSGFLVYSD 235
```

```
RESULT 3
COT1 HUMAN STANDARD; PRT; 281 AA.
ID COT1 HUMAN
AC Q9BXJ1; Q96NF2; Q9GZR4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement-c1q tumor necrosis factor-related protein 1 precursor
DE (G protein coupled receptor interacting protein) (GIP).
GN C1QTNF1 OR CTRP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Innamorati G., Le Gouill C., Whang I., Birnbaumer M.;
RT "GIP, a putative GPCR interacting protein.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Sheppard P.O.;
RT "Homo sapiens complement-c1q tumor necrosis factor-related protein.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedln T.B., Toshlyuk S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [4]
RP SEQUENCE OF 83-281 FROM N.A.
RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA Hotuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J.,
RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C1Q domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AJ272138; CAC20425.1; ALT_INIT.
```

DR EMBL; AF232905; AAG44303.1; ALT_INIT.
DR EMBL; AF329840; AAK17964.1; .
DR EMBL; BC021553; AAH21553.1; .
DR EMBL; AK055541; BAB70947.1; .
DR Genew; HGNC:14324; C1QTNF1.
DR InterPro; IPR001073; C1Q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; FALSE_NEG.
KW Collagen; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 281 COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT DOMAIN 99 140 RELATED PROTEIN 1.
FT DOMAIN 147 281 COLLAGEN-LIKE.
FT CONFLICT 22 22 C1Q.
FT CONFLICT 241 241 L -> P (IN REF. 1).
FT CONFLICT 241 241 R -> Q (IN REF. 1).
SQ SEQUENCE 281 AA; 31743 MW; 49E248CB88ACFB7C CRC64;

Query Match 22.7%; Score 202.5; DB 1; Length 281;
Best Local Similarity 35.5%; Pred. No. 7.4e-12;
Matches 60; Conservative 18; Mismatches 58; Indels 33; Gaps 6;

QY 5 APARGPPAP-----PEPR-----SAFSAARTSLVSGSDAGPGRH-----QPLAFD 45
DB 119 AGARGHTGPKGQKSGMGAPGERCKSHYAFAVSGRKK-----PMHSNHYQYTFD 168
QY 46 TEFVNIIGDFFDAAGVFRCLPAGVFFSFTLGKLPKRTLSVKLMKNRDEVQAMTYDDGAS 105
DB 169 TEFVNIYDHFMFTGKFCYCPGLYFFSLNVHTMNOKETYLHIMKNEBEVILFAQYG-- 226
QY 106 RRREMOSQSVMLALRRGDVWMLSHDHDGYGAYSNH--GKYITFSGFLV 152
DB 227 DRSIMOSQSLMLELRQDQVWVRLYKGERENAIFFSELDITYITFSGYLV 275

RESULT 4
COT6 HUMAN STANDARD; PRT; 278 AA.
ID COT6_HUMAN
AC Q9BXI9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement-c1q tumor necrosis factor-related protein 6 precursor.
GN C1QTNF6 OR CTRP6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Piddington C.S., Sheppard P.O.;
RT "Homo sapiens complement-c1q tumor necrosis factor-related protein.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- SIMILARITY: Contains 1 collagenous domain.
CC -!- SIMILARITY: Contains 1 C1Q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AF329842; AAK17966.1; .
DR EMBL; BC020551; AAH20551.1; .
DR Genew; HGNC:14343; C1QTNF6.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR PROSITE; PS01113; C1Q; FALSE_NEG.
KW Collagen; Signal.
FT SIGNAL 1 46
FT CHAIN 47 278 POTENTIAL.
FT DOMAIN 97 138 COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT DOMAIN 139 278 RELATED PROTEIN 6.
FT CARBOHYD 91 91 COLLAGEN-LIKE.
FT CONFLICT 21 21 C1Q.
FT CONFLICT 21 21 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 278 AA; 30861 MW; 27A82CA863F23D47 CRC64;

Query Match 21.7%; Score 194; DB 1; Length 278;
Best Local Similarity 41.8%; Pred. No. 4.6e-11;
Matches 61; Conservative 10; Mismatches 65; Indels 10; Gaps 5;

QY 10 PPAPPEPR-SAFSAARTSLVSGSDAGPGRHQPPLAFDTEFVNIIGDFFDAAGVFRCLRP 68
DB 135 PGAPCQKRFPAFVSGRKTA-----HSGEDQTLFPRVFNLDGCFMATGQFAAPLRG 189
QY 69 AYFFSFTLGKLPKRTLSVKLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRRGDVWML 128
DB 190 IYFSLNVHSMWYKETYVHIMNOKE--AVILYAQPSERSIMOSQSVMLDLAYGDRVWVR 247
QY 129 SHDHDGYGAYSNH-GKYITFSGFLV 152
DB 248 LFKRQRENAIYNSNDFDYITFSGHLI 273

RESULT 5
C1RF MOUSE STANDARD; PRT; 258 AA.
ID C1RF_MOUSE
AC O88992;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C1q-related factor precursor.
GN C1QRF OR CRF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99097006; PubMed=9878755;
RA Berube N.G., Swanson X.H., Bertram M.J., Kittle J.D., Didenko V.,
RA Baekin D.S., Smith J.R., Pereira-Smith O.M.;
RT "Cloning and characterization of CRF, a novel C1q-related factor,

```
RT expressed in areas of the brain involved in motor function.";
RL Brain Res. Mol. Brain Res. 63:233-240(1999).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAINSTEM. MORE ABUNDANT IN AREAS
CC OF THE NERVOUS SYSTEM INVOLVED IN MOTOR FUNCTION, SUCH AS THE
CC PURKINJE CELLS OF THE CEREBELLUM, THE ACCESSORY OLIVARY NUCLEUS,
CC THE PONS AND THE RED NUCLEUS.
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C1Q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF095155; AAC64187.1; -.
DR MGI; MGI:1344400; C1qrf.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Collagen; Signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 258 C1Q-RELATED FACTOR.
FT DOMAIN 67 115 COLLAGEN-LIKE.
FT DOMAIN 123 258 C1Q.
SQ SEQUENCE 258 AA; 26485 MW; F776E2D206EBF763 CRC64;

Query Match 21.5%; Score 192.5; DB 1; Length 258;
Best Local Similarity 29.7%; Pred. No. 5.8e-11;
Matches 55; Conservative 20; Mismatches 67; Indels 43; Gaps 5;

QY 9 GPPAPPEPSAFAARTSLVSGDAGPGRHQP----- 41
   |||||
DB 79 GPPGPPGDRGPPGPPVPPGPEKEGPPGPPGLPGSSGSAISTATYTTVPVAFYAGLKN 138
   |||||

QY 42 -----LAFDTEFVNIGDGFDAAGVPRCLPGAYFFSFTLGKLP---KTLVSKLMK 90
   |||||
DB 139 PHEGYEVLFKFDVVTNLGNNDYDASGKFTCNIPGTYFTYHV--LMRGDGTSMWADLCK 196
   |||||

QY 91 NRDEVOAMTYDDGASRRREMOSQSVMLALRGDAVWLLSHDHGAYGAYSNHGKYITFSGF 150
   |||||
DB 197 N-GQVRASAIADADQNYDYASNSVILHLDAGDEVFIKLDGKAGHGNSN--KYSTFSGF 253
   |||||

QY 151 LVYPD 155
   |||||
DB 254 IIVSD 258

RESULT 6
C1RF HUMAN STANDARD; PRT; 258 AA.
ID C1RF HUMAN
AC 075973;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE C1q-related factor precursor.
DE C1q-related factor precursor.
GN C1QRF OR CRF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99097006; PubMed=9878755;
RA Berube N.G., Swanson X.H., Bertram M.J., Kittle J.D., Didenko V.,
RA Baekin D.S., Smith J.R., Pereira-Smith O.M.;
```

```
RT "Cloning and characterization of CRF, a novel C1q-related factor,
RT expressed in areas of the brain involved in motor function.";
RL Brain Res. Mol. Brain Res. 63:233-240(1999).
CC [2]
CC SEQUENCE FROM N.A.
CC Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
CC Nickerson D.A.;
CC Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
CC [3]
CC SEQUENCE FROM N.A.
CC TISSUE=Placenta;
CC MEDLINE=22388257; PubMed=12477932;
CC Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
CC Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
CC Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
CC Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
CC Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
CC Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
CC Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
CC Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
CC Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
CC Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
CC Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
CC Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
CC Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
CC Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
CC Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
CC Butterfield Y.S.N., Krzyzanski M.I., Skalska U., Smailus D.E.,
CC Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
CC "Generation and initial analysis of more than 15,000 full-length
CC human and mouse cDNA sequences.";
CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAINSTEM.
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C1Q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF095154; AAC64186.1; -.
DR EMBL; AF410771; AAK95248.1; -.
DR EMBL; BC008798; AAH08798.1; -.
DR GO; GO:0007626; P:locomotory behavior; NAS.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Collagen; Signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 258 C1Q-RELATED FACTOR.
FT DOMAIN 67 115 COLLAGEN-LIKE.
FT DOMAIN 123 258 C1Q.
SQ SEQUENCE 258 AA; 26452 MW; 52C51CDF59CAE2BF CRC64;

Query Match 21.2%; Score 189.5; DB 1; Length 258;
Best Local Similarity 32.0%; Pred. No. 1.1e-10;
Matches 54; Conservative 22; Mismatches 58; Indels 35; Gaps 6;

QY 9 GPPAP-----EPRSAFSAARTSLVSGDAGPGRHOPLAFTTEFVN 50
   |||||
DB 103 GKPPGRLPGAGSGAISTATYTTVPVAFYAGL-----KNPHEGYEVLFKFDVVTN 154
   |||||

QY 51 IGGDFDAAGVFRCLPGAYFFSFTLGKLP---KTLVSKLMKNRDEVOAMTYDDGASR 106
   |||||
DB 155 LGNNYDASGKFTCNIPGTYFTYHV--LMRGDGTSMWADLCKN-GQVRASAIADADQ 211
   |||||
```


OY 107 RREMOSQSVMLALRRGDAVWMLSHDHGCGAYSNHGKYTEGSEFLVYPD 155
Db 212 NYDYASNSVILHLHDAGDEVFIKLDGKRAHGNSN--KYSTPSGFIITYSD 258

RESULT 7

APM1_MOUSE STANDARD; PRT; 247 AA.
ID APM1_MOUSE Q62400; Q9DC68;
AC Q60994; Q62400; Q9DC68;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Adiponectin precursor (30 kDa adipocyte complement-related protein)
DE (ACRP30) (Adipocyte specific protein AdipoQ).
GN APM1 OR ACRP30 OR ADIPOQ.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adipocyte;
RX MEDLINE=96070757; PubMed=7592907;
RA Scherer P.E., Williams S., Fogliano M., Baldini G., Lodish H.F.;
RT "A novel serum protein similar to C1q, produced exclusively in
adipocytes."
RL J. Biol. Chem. 270:26746-26749(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fibroblast;
RX MEDLINE=96209999; PubMed=8631877;
RA Hu E., Liang P., Spiegelman B.M.;
RT "AdipoQ is a novel adipose-specific gene dysregulated in obesity."
RL J. Biol. Chem. 271:10697-10703(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX PubMed=11162643;
RA Das K., Lin Y., Widen E., Zhang Y., Scherer P.E.;
RT "Chromosomal localization, expression pattern, and promoter analysis
of the mouse gene encoding adipocyte-specific secretory protein
Acrp30."
RL Biochem. Biophys. Res. Commun. 280:1120-1129(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,
Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
Schrml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Haegawa Y., Kawai H., Kohetsuki S.,
Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [5]
RP FUNCTION.
RX MEDLINE=21372498; PubMed=11479627;
RA Yamauchi T., Kamon J., Waki H., Teranuchi Y., Kubota N., Hara K.,
Mori Y., Ide T., Murakami K., Tsuboyama-Kasaoka N., Ezaki O.,
Akanuma Y., Gavrilova O., Vinson C., Reitman M.L., Kagechika H.,

RA Shudo K., Yoda M., Nakano Y., Tobe K., Nagai R., Kimura S., Tomita M.,
RA Froguel P., Kadowaki T.;
RT "The fat-derived hormone adiponectin reverses insulin resistance
RT associated with both lipodystrophy and obesity.";
RL Nat. Med. 7:941-946(2001).
RN [6]
RP FUNCTION.
RX MEDLINE=21372499; PubMed=11479628;
RA Berg A.H., Combs T.P., Du X., Brownlee M., Scherer P.E.;
RT "The adipocyte-secreted protein Acrp30 enhances hepatic insulin
action.";
RL Nat. Med. 7:947-953(2001).
CC -!- FUNCTION: IMPORTANT NEGATIVE REGULATOR IN HEMATOPOIESIS AND IMMUNE
CC SYSTEMS; MAY BE INVOLVED IN ENDING INFLAMMATORY RESPONSES THROUGH
CC ITS INHIBITORY FUNCTIONS. INHIBITS ENDOTHELIAL NF-KAPPA-B SIGNALING
CC THROUGH A CAMP-DEPENDENT PATHWAY. INHIBITS TNF-ALPHA-INDUCED
CC EXPRESSION OF ENDOTHELIAL ADHESION MOLECULES. INVOLVED IN THE
CC CONTROL OF FAT METABOLISM AND INSULIN SENSITIVITY.
CC -!- SUBUNIT: Homooligomer.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Synthesized exclusively by adipocytes and
CC secreted into plasma.
CC -!- INDUCTION: DURING HORMONE-INDUCED ADIPOSE DIFFERENTIATION AND
CC ACTIVATED BY INSULIN.
CC -!- SIMILARITY: Contains 1 collagenous domain.
CC -!- SIMILARITY: Contains 1 C1Q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U37222; AAA80543.1; -.
DR EMBL; U49915; AAB06706.1; -.
DR EMBL; AF304466; AAK13417.1; -.
DR EMBL; AK003138; BAB22597.1; -.
DR PDB; 1C28; 07-SEP-99.
DR MGD; MGI:106675; Acrp30.
DR GO; GO:0005576; C:extracellular; IDA.
DR GO; GO:0005515; F:protein binding activity; IPI.
DR GO; GO:0006635; P:fatty acid beta-oxidation; IMP.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR ProDom; PD000007; C1q_helix; 1.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Hormone; Collagen; Signal; Repeat; Hydroxylation; Plasma;
KW Polymorphism; 3D-structure.
FT SIGNAL 1 17
FT CHAIN 18 247
FT DOMAIN 45 110
FT DOMAIN 111 247
FT DISULFID 39 39
FT MOD_RES 47 47
FT MOD_RES 50 50
FT MOD_RES 56 56
FT MOD_RES 65 65
FT MOD_RES 79 79
FT MOD_RES 98 98
FT MOD_RES 107 107
FT VARIANT 113 113
FT CONFLICT 50 50
FT CONFLICT 74 74
FT CONFLICT 117 117
FT CONFLICT 148 148
FT CONFLICT 243 243
SQ SEQUENCE 247 AA; 26841 MW; 137B687D873988C4 CRC64;

[illegible]

```

DR   Genew; HGNC:14325; C1QTNF2.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 2.
DR   PRINTS; PR00007; COMPLEMENTC1Q.
DR   SMART; SM00110; C1Q; 1.
DR   PROSITE; PS01113; C1Q; 1.
DR   KX   Collagen; Signal.
FT   SIGNAL      1      15      POTENTIAL.
FT   CHAIN       16      285     COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT                                     RELATED PROTEIN 2.
FT   DOMAIN      40      141     COLLAGEN-LIKE.
FT   DOMAIN      143     285     C1Q.
SQ   SEQUENCE    285 AA; 29952 MW; 7E31FF9868D4EDFA CRC64;
Query Match      21.0%; Score 187.5; DB 1; Length 285;
Best Local Similarity 32.5%; Pred. No. 1.9e-10;
Matches 52; Conservative 31; Mismatches 52; Indels 25; Gaps 8;

QY   8  RGPAPPEP-----RSAFSAARTSLVGS DADGPPRHQ-PLAFDFEVNIGDFDAA 58
Db    132 KGEPLPGPCSCGSGHTKSAFSAVATKSY-----PRERLPKIPDKILMEGGHYNAS 183
QY   59 AGVFCRLPGAYFFESFTLGKLPKRTLSVKMKNRDEVQAMTYDDGASRRREMOSQSVMLA 118
Db    184 SGKFCVCGVPGIYYFTYDI-TLANKHLAIGLVHN-GQYRIRTFDANTG-NHDVASGSTILA 240
QY   119 LRRGDAVWLLSHDHDGYGAYSNHGKYIT--FSGFLVYPD 155
Db    241 LKQGDVWLQIFYSQNGLF--YDPYWTDSLFTGFLIYAD 278

RESULT 9
GLIC_MOUSE
ID   GLIC_MOUSE      STANDARD;          PRT;          255 AA.
AC   Q9ESN4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Glacolin precursor (C1q-like protein).
GN   C1QL.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20428709; PubMed=10862616;
RA   Koide T., Aso A., Yorihuzi T., Nagata K.;
RT   "Conformational requirements of collagenous peptides for recognition
RT   by the chaperrone protein HSP47."
RL   J. Biol. Chem. 275:27957-27963(2000).
CC   -1- TISSUE SPECIFICITY: EXPRESSED IN GLIAL CELLS.
CC   -1- SIMILARITY: Contains 1 collagenous domain.
CC   -1- SIMILARITY: Contains 1 C1q domain.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; AB044560; BAB15806.1; -.
DR   MGD; MGI:2387350; C1QL.
DR   GO; GO:0005515; F:protein binding activity; IPI.
DR   InterPro; IPR001073; C1q.
DR   InterPro; IPR000087; Collagen.
DR   Pfam; PF00386; C1q; 1.
DR   Pfam; PF01391; Collagen; 1.
DR   PRINTS; PR00007; COMPLEMENTC1Q.

```

```
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Collagen; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 255 GLIACOLIN.
FT DOMAIN 61 111 COLLAGEN-LIKE.
FT DOMAIN 120 255 C1Q.
SQ SEQUENCE 255 AA; 26687 MW; 529FBAF4B2191BC1 CRC64;

Query Match 20.7%; Score 185.5; DB 1; Length 255;
Best Local Similarity 28.0%; Pred. No. 2.6e-10;
Matches 53; Conservative 21; Mismatches 72; Indels 43; Gaps 5;

QY 5 APARGPAPPEPRSAFSAARTSLVGSADGPPRHQP----- 41
Db 72 AGPRGPPEGPPEPRGVPGEKGEPRGGLPGRPGAPGLNAGAI SATYSTVPKIAFYA 131
QY 42 -----LAFDTEFVNIGGDFDAAGVFRCLPGAYFFSFTLGKLPB---KTLSV 86
Db 132 GLKRQHEGYEVLFKEDDVVTNLGNHYDPTTEKFTCSIPGIYFTYHV--LMRGDGTSMWA 189
QY 87 KLMKNRDEVQAMITYDDGASRRREMOSQVMALRRGDAVWLISHDHGAYGAYSNHGKYIT 146
Db 190 DLCKN-NQVRASAIQADADQNYDASNSVVLHLEPGDEVYIKLDGKAHG--GNNNKYST 246
QY 147 FSGFLVYPD 155
Db 247 FSGFTIYAD 255

RESULT 10
C1QB_HUMAN STANDARD; PRT; 251 AA.
AC P02746; Q96H17;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement C1q subcomponent, B chain precursor.
GN C1QB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86076906; Pubmed=3000358;
RA Reid K.B.M.;
RT "Molecular cloning and characterization of the complementary DNA and
RT gene coding for the B-chain of subcomponent C1q of the human
RT complement system.";
RL Biochem. J. 231:729-735(1985).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; Pubmed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshlyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
```

```
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 26-133.
RX MEDLINE=80020137; Pubmed=486087;
RA Reid K.B.M.;
RT "Complete amino acid sequences of the three collagen-like regions
RT present in subcomponent C1q of the first component of human
RT complement.";
RL Biochem. J. 179:367-371(1979).
RN [4]
RP SEQUENCE OF 26-193.
RX MEDLINE=79041552; Pubmed=708376;
RA Reid K.B.M., Thompson E.O.P.;
RT "Amino acid sequence of the N-terminal 108 amino acid residues of the
RT B chain of subcomponent C1q of the first component of human
RT complement.";
RL Biochem. J. 173:863-868(1978).
RN [5]
RP SEQUENCE OF 134-251.
RX MEDLINE=82283890; Pubmed=6981411;
RA Reid K.B.M., Gagnon J., Frampton J.;
RT "Completion of the amino acid sequences of the A and B chains of
RT subcomponent C1q of the first component of human complement.";
RL Biochem. J. 203:559-569(1982).
RN [6]
RP SEQUENCE OF 224-251 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=85038855; Pubmed=6208566;
RA Reid K.B.M., Bentley D.R., Wood K.J.;
RT "Cloning and characterization of the complementary DNA for the B
RT chain of normal human serum C1q.";
RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 306:345-354(1984).
RN [7]
RP REVIEW OF C1Q DEFICIENCY.
RX MEDLINE=98450587; Pubmed=9777412;
RA Petry F.;
RT "Molecular basis of hereditary C1q deficiency.";
RL Immunobiology 199:286-294(1998).
CC -!- FUNCTION: C1Q ASSOCIATES WITH THE PROENZYMES C1R AND C1S TO YIELD
CC C1, THE FIRST COMPONENT OF THE SERUM COMPLEMENT SYSTEM. THE
CC COLLAGEN-LIKE REGIONS OF C1Q INTERACT WITH THE CA(2+)-DEPENDENT
CC C1R(2)C1S(2) PROENZYME COMPLEX, AND EFFICIENT ACTIVATION OF C1
CC TAKES PLACE ON INTERACTION OF THE GLOBULAR HEADS OF C1Q WITH THE
CC FC REGIONS OF IGG OR IGM ANTIBODY PRESENT IN IMMUNE COMPLEXES.
CC -!- SUBUNIT: C1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q, R
CC AND S IN THE MOLAR RATION OF 1:2:2. C1Q SUBCOMPONENT IS COMPOSED
CC OF NINE SUBUNITS, SIX OF WHICH ARE DISULFIDE-LINKED DIMERS OF THE
CC A AND B CHAINS, AND THREE OF WHICH ARE DISULFIDE-LINKED DIMERS OF
CC THE C CHAIN.
CC -!- PTM: O-linked glycans consist of Glc-Gal disaccharides bound to
CC the oxygen atom of post-translationally added hydroxyl groups.
CC -!- DISEASE: Defects in C1QB are a cause of C1q deficiency
CC [MIM:120570]. It is a rare genetic disorder which is associated
CC with recurrent infections and a high prevalence of lupus
CC erythematosus-like symptoms. It is characterized by a loss of
CC activation of the complement classical pathway.
CC -!- SIMILARITY: Contains 1 collagenous domain.
CC -!- SIMILARITY: Contains 1 C1q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X03084; CAA26880.1; -
CC EMBL; BC008983; AAH08983.1; ALT_INIT.
CC EMBL; M36278; AAC41692.1; -
CC Genew; HGNC:1242; C1QB.
CC MIM; 120570; -.
```

DR GO; GO:0005602; C:complement component C1q complex; TAS.
DR GO; GO:0003811; F:complement activity activity; TAS.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Complement pathway; Plasma; Hydroxylation; Glycoprotein; Collagen;
Repeat; Signal; Disease mutation; Pyroglutamate carboxylic acid.
FT SIGNAL 1 25
FT CHAIN 26 251
FT DOMAIN 29 112
FT DOMAIN 113 251
FT MOD_RES 26 26
FT MOD_RES 29 29
FT MOD_RES 33 33
FT MOD_RES 36 36
FT MOD_RES 39 39
FT MOD_RES 42 42
FT MOD_RES 51 51
FT MOD_RES 54 54
FT MOD_RES 57 57
FT CARBOHYD 57 57
FT MOD_RES 60 60
FT CARBOHYD 60 60
FT MOD_RES 63 63
FT MOD_RES 75 75
FT MOD_RES 81 81
FT MOD_RES 84 84
FT MOD_RES 90 90
FT MOD_RES 96 96
FT CARBOHYD 96 96
FT MOD_RES 99 99
FT MOD_RES 102 102
FT MOD_RES 105 105
FT MOD_RES 108 108
FT CARBOHYD 108 108
FT VARIANT 40 40
FT CONFLICT 26 26
FT CONFLICT 83 83
FT CONFLICT 98 98
SQ SEQUENCE 251 AA; 26459 MW; 78C5752E267A0EF7 CRC64;

Query Match 20.4%; Score 182; DB 1; Length 251;
Best Local Similarity 29.1%; Pred. No. 5.5e-10;
Matches 48; Conservative 32; Mismatches 61; Indels 24; Gaps 5;

OY 6 PARGPAPPP-----RSARFSAARTSLVSGDAGPGRHQPLAFDTEFNIG 52
DB 95 PKGGPGAPGAPGPKGESGDKATQKIAFSATRTIN-----PLRRDQITIRFDHVTNN 148
OY 53 GDFDAAGVFRCLPGAYFFSFTLGKLPKRLSVKLMKNRDEVQAMI-YDDGASRRREM 111
DB 149 NNYEPRSGKFTCKVGLYFTYHASS--RGNLCVNLMRGRERAQKVTFCDYAYNTFOVT 206
OY 112 SOSVMLALRGDAVWLLSHDHDGYGAYSNHGKYITFSGFLVYDPL 156
DB 207 TGMVNLKLEGENVFLQATDKN--SLIGMEGANSIFSGLFLPDM 249

RESULT 11
HP25_TAMSI
ID HP25_TAMSI STANDARD; PRT; 215 AA.
AC Q06576;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hibernation-associated plasma protein HP-25 precursor (hibernator-specific blood complex, 25 kDa subunit).
OS Tamias sibiricus (Siberian chipmunk) (Asian chipmunk).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Sciurinae;
OC Tamias.
OX NCBI_TaxID=64680;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93180798; PubMed=8441393;
RA Takamatsu N., Ohba K., Kondo J., Kondo N., Shiba T.;
RT "Hibernation-associated gene regulation of plasma proteins with a
RT collagen-like domain in mammalian hibernators.";
RL Mol. Cell. Biol. 13:1516-1521(1993).
RN [2]
RP SEQUENCE OF 29-62; 84-130; 172-183; 187-192 AND 201-215.
RC TISSUE=Plasma;
RX MEDLINE=92112696; PubMed=1730610;
RA Kondo N., Kondo J.;
RT "Identification of novel blood proteins specific for mammalian
RT hibernation.";
RL J. Biol. Chem. 267:473-478(1992).
CC -1- FUNCTION: PLASMA PROTEINS HP-20, HP-25, HP-27 AND HP-55 FORM A
CC 140 kDa COMPLEX VIA DISULFIDE BONDS IN THE PLASMA AND ARE
CC HIBERNATION SPECIFIC.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
CC -1- DEVELOPMENTAL STAGE: THE PROTEIN COMPLEX DISAPPEARS FROM THE
CC PLASMA AT ONSET OF HIBERNATION AND REAPPEARS AS HIBERNATION
CC CEASES.
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C1Q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D12975; BAA02352.1; -
DR PIR; B48150; B48150.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Signal; Collagen; Glycoprotein; Plasma; Multigene family.
FT SIGNAL 1 28
FT CHAIN 29 215
FT DOMAIN 40 81
FT DOMAIN 83 215
FT CARBOHYD 167 167
SQ SEQUENCE 215 AA; 22664 MW; AFE03206917EA530 CRC64;

Query Match 19.9%; Score 177.5; DB 1; Length 215;
Best Local Similarity 32.4%; Pred. No. 1.2e-09;
Matches 58; Conservative 21; Mismatches 49; Indels 51; Gaps 10;

OY 5 AP-ARGPAPPP-----EPRSAFSAARTSLVSGDAGPGRHQPLA 43
DB 56 APGALGPPGPPGVPPIPGPPGPDVEKCSSRPKSAFAVKL-----SERPPEP-FQPIV 108
OY 44 FDTFVNIGDFAAGVFRCLPGAYFFSFTLGKLPKRLSVKLMKNRDEVQAMYYDDG 103
DB 109 FKEALYNQEGHFNMATGEFSCVLPGVNFGFDI-RLFGSSVKIRLM--RDGIQV----- 159
OY 104 ASRRREMOSQ-----SVMLALRGDAVWLLSHDHDGYGAYSNHG-KYITFSGFLVY 153
DB 160 --REKXAQANDSYKHAMGSYVMAIGKGDVWLESKLR--GTESEKGIHIVFGYLLV 213

RESULT 12
APM1 HUMAN STANDARD; PRT; 244 AA.
ID APM1_HUMAN
AC Q15848;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Adiponectin precursor (30 kDa adipocyte complement-related protein)
DE (ACRP30) (Adipose most abundant gene transcript 1) (apm-1) (Gelatin-
binding protein).
GN APM1 OR ACRP30 OR GBP28.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adipose tissue;
RX MEDLINE=96224171; PubMed=8619847;
RA Maeda K., Okubo K., Shimomura I., Funahashi T., Matsuzawa Y.,
RA Matsubara K.;
RT "cDNA cloning and expression of a novel adipose specific collagen-like
RT factor, apm1 (Adipose Most abundant Gene transcript 1).";
RL Biochem. Biophys. Res. Commun. 221:286-289(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99196984; PubMed=10095105;
RA Saito K., Tobe T., Minoshima S., Asakawa S., Sumiya J., Yoda M.,
RA Nakano Y., Shimizu N., Tomita M.;
RT "Organization of the gene for gelatin-binding protein (GBP28).";
RL Gene 229:67-73(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99333693; PubMed=10403784;
RA Schaeffler A., Orso E., Palitzsch K.D., Buechler C., Drobnik W.,
RA Fuerst A., Scholmerich J., Schmitz G.;
RT "The human apm-1, an adipocyte-specific gene linked to the family of
RT TNF's and to genes expressed in activated T cells, is mapped to
RT chromosome 1q21.3-q23, a susceptibility locus identified for familial
RT combined hyperlipidemia (FCH).";
RL Biochem. Biophys. Res. Commun. 260:416-425(1999).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=20417747; PubMed=10961870;
RA Yokota T., Oritani K., Takahashi I., Ishikawa J., Matsuyama A.,
RA Ouchi N., Kihara S., Funahashi T., Tenner A.J., Tomiyama Y.,
RA Matsuzawa Y.;
RT "Adiponectin, a new member of the family of soluble defense collagens,
RT negatively regulates the growth of myelomonocytic progenitors and the
RT functions of macrophages.";
RL Blood 96:1723-1732(2000).
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=20440368; PubMed=10982546;
RA Ouchi N., Kihara S., Arita Y., Okamoto Y., Maeda K., Kuriyama H.,
RA Hotta K., Nishida M., Takahashi M., Muraguchi M., Ohmoto Y.,
RA Nakamura T., Yamashita S., Funahashi T., Matsuzawa Y.;
RT "Adiponectin, an adipocyte-derived plasma protein, inhibits
RT endothelial NF-kappaB signaling through a CAMP-dependent pathway.";
RL Circulation 102:1296-1301(2000).
RN [6]
RP FUNCTION.
RX MEDLINE=21372498; PubMed=11479627;
RA Yamauchi T., Kamon J., Waki H., Teranuchi Y., Kubota N., Hara K.,
RA Mori Y., Ide T., Murakami K., Tsuboyama-Kasaoka N., Ezaki O.,
RA Akanuma Y., Gavrilova O., Vinson C., Reitman M.L., Kagechika H.,
RA Shudo K., Yoda M., Nakano Y., Tobe K., Nagai R., Kimura S., Tomita M.,
RA Froguel P., Kadowaki T.;
RT "The fat-derived hormone adiponectin reverses insulin resistance
RT associated with both lipotrophy and obesity.";
RL Nat. Med. 7:941-946(2001).
RN [7]

RP VARIANT ADIPONECTIN DEFICIENCY CYS-112.
RX MEDLINE=20378830; PubMed=10918532;
RA Takahashi M., Arita Y., Yamagata K., Matsukawa Y., Okutomi K.,
RA Horie M., Shimomura I., Hotta K., Kuriyama H., Kihara S., Nakamura T.,
RA Yamashita S., Funahashi T., Matsuzawa Y.;
RT "Genomic structure and mutations in adipose-specific gene,
RT adiponectin.";
RL Int. J. Obes. Relat. Metab. Disord. 24:861-868(2000).
RN [8]
RP VARIANTS ARG-84; MET-117; THR-164; SER-221 AND PRO-241.
RX MEDLINE=21671103; PubMed=11812766;
RA Hara K., Boutin P., Mori Y., Tobe K., Dina C., Yasuda K., Yamauchi T.,
RA Otobe S., Okada T., Eto K., Kadowaki H., Hagura R., Akanuma Y.,
RA Yarakai Y., Nagai R., Taniyama M., Matsubara K., Yoda M., Nakano Y.,
RA Kimura S., Tomita M., Kimura S., Ito C., Froguel P., Kadowaki T.;
RT "Genetic variation in the gene encoding adiponectin is associated with
RT an increased risk of type 2 diabetes in the Japanese population.";
RL Diabetes 51:536-540(2002).
CC -I- FUNCTION: IMPORTANT NEGATIVE REGULATOR IN HEMATOPOIESIS AND IMMUNE
CC SYSTEMS; MAY BE INVOLVED IN ENDING INFLAMMATORY RESPONSES THROUGH
CC ITS INHIBITORY FUNCTIONS. INHIBITS ENDOTHELIAL NF-KAPPAB SIGNALING
CC THROUGH A CAMP-DEPENDENT PATHWAY. INHIBITS TNF-ALPHA-INDUCED
CC EXPRESSION OF ENDOTHELIAL ADHESION MOLECULES. INVOLVED IN THE
CC CONTROL OF FAT METABOLISM AND INSULIN SENSITIVITY.
CC -I- SUBUNIT: HOMODIGOMER (POTENTIAL).
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- TISSUE SPECIFICITY: Synthesized exclusively by adipocytes and
CC secreted into plasma.
CC -I- DISEASE: Defects in APM1 are the cause of adiponectin deficiency
CC [MIM:605441], resulting in very low concentration of plasma
CC adiponectin. Decreased adiponectin plasma levels are associated
CC with obesity insulin resistance, and diabetes type 2.
CC -I- PHARMACEUTICAL: Adiponectin might be used in the treatment of
CC diabetes type 2 and insulin resistance.
CC -I- SIMILARITY: Contains 1 collagenous domain.
CC -I- SIMILARITY: Contains 1 C1Q domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; D45371; BAA08227.1; --
DR EMBL; AB012165; BAA86716.1; --
DR EMBL; AB012164; BAA86716.1; JOINED.
DR EMBL; AJ131460; CAB52413.1; --
DR EMBL; AJ131461; CAB52413.1; JOINED.
DR PIR; JC4708; JC4708.
DR MIM; 605441; --
DR GO; GO:0006091; P:energy pathways; TAS.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR ProDom; PD000007; C1q_helix; 1.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Hormone; Collagen; Signal; Repeat; Hydroxylation; Plasma;
KW Polymorphism; Disease mutation; Obesity; Diabetes mellitus.
FT SIGNAL 1 14
FT CHAIN 15 244 ADIPONECTIN.
FT DOMAIN 42 107 COLLAGEN-LIKE.
FT DOMAIN 108 244 C1Q.
FT DISULFID 36 36 INTERCHAIN (BY SIMILARITY).
FT MOD_RES 44 44 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 47 47 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 53 53 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 62 62 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 71 71 HYDROXYLATION (BY SIMILARITY).


```

FT  MOD_RES      76      76      HYDROXYLATION (BY SIMILARITY).
FT  MOD_RES      86      86      HYDROXYLATION (BY SIMILARITY).
FT  MOD_RES      95      95      HYDROXYLATION (BY SIMILARITY).
FT  MOD_RES     104     104      HYDROXYLATION (BY SIMILARITY).
FT  MOD_RES      84      84      G -> R.
FT  MOD_RES      112     112      /FTID=VAR_013273.
FT  MOD_RES      112     112      R -> C (In adiponectin deficiency).
FT  MOD_RES      117     117      /FTID=VAR_013274.
FT  MOD_RES      117     117      V -> M.
FT  MOD_RES      164     164      /FTID=VAR_013275.
FT  MOD_RES      164     164      I -> T.
FT  MOD_RES      221     221      /FTID=VAR_013276.
FT  MOD_RES      221     221      R -> S.
FT  MOD_RES      241     241      /FTID=VAR_013277.
FT  MOD_RES      241     241      H -> P.
FT  MOD_RES      241     241      /FTID=VAR_013278.
SQ  SEQUENCE     244 AA; 26414 MW; 64D8C6C1204B1018 CRC64;

Query Match      19.6%; Score 175; DB 1; Length 244;
Best Local Similarity 31.1%; Pred. No. 2.4e-09;
Matches 51; Conservative 24; Mismatches 65; Indels 24; Gaps 6;

QY  7 ARGPAPP-----EP-----RSAFSAARTSLVGSADGPRHQPLAFDTEFNIG 52
    |||||
DB  88 AEGPRGPFPIQGRKGEPEGAGVYVRSASF-----VGLFTYVTIPNMPIRFKIFYNQO 140
    |||||

QY  53 GDFDAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMTYDDGASRRREMOS 112
    |||||
DB  141 NHYDGSTGKFHCNIPGLYFYAYHI-TVMKDKVSLFK-KDKAMLFTYDQYQENNVQAS 198
    |||||

QY  113 QSVMLALRRGDAVWLSH-DHDGYGAYSNGKTYTPSGFLVYPD 155
    |||||
DB  199 GSVLLHLLEVGDQVWLQVYGEGERNGLYADNDNDSTFTGFLLYHD 242
    |||||

RESULT 13
COLE_LEPMA
ID  COLE_LEPMA      STANDARD;      PRT;      419 AA.
AC  P98085; Q91080;
DT  01-NOV-1995 (Rel. 32, Created)
DT  01-NOV-1995 (Rel. 32, Last sequence update)
DT  15-JUL-1998 (Rel. 36, Last annotation update)
DE  Inner ear-specific collagen precursor (Saccular collagen).
OS  Lepomis macrochirus (Bluegill).
OC  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC  Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC  Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC  Centrarchidae; Lepomis.
OC  NCBI_Taxid=13106;
OX  NCBI_Taxid=13106;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=95167486; PubMed=7863331;
RA  Davis J.G., Oberholtzer J.C., Burns F.R., Greene M.I.;
RT  "Molecular cloning and characterization of an inner ear-specific
    structural protein."
RL  Science 267:1031-1034(1995).
RN  [2]
RP  CONCEPTUAL TRANSLATION.
RA  Gibson T.;
RL  Submitted (MAR-1995) to the SWISS-PROT data bank.
CC  -1- FUNCTION: FORMS A MICROSTRUCTURAL MATRIX WITHIN THE OTOLITHIC
    MEMBRANE (PROBABLY).
CC  -1- TISSUE SPECIFICITY: SPECIALIZED SECRETORY SUPPORTING CELLS AT THE
    OUTER PERIMETER OF THE SACCULAR EPITHELIUM.
CC  -1- SIMILARITY: Contains 1 C1q domain.
CC  -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A FRAMESHIFT HAD TO BE
    INTRODUCED FOR POSITIONS 391-419 SO AS TO MAXIMIZE THE SIMILARITY
    WITH OTHER SHORT-CHAIN COLLAGENS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
    between the Swiss Institute of Bioinformatics and the EMBL outstation -
    the European Bioinformatics Institute. There are no restrictions on its
    use by non-profit institutions as long as its content is in no way

```

```

CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; U17431; AAA69978.1; ALT_FRAME.
DR  InterPro; IPR001073; C1q.
DR  InterPro; IPR000087; Collagen.
DR  Pfam; PF00386; C1q; 1.
DR  Pfam; PF01391; Collagen; 3.
DR  PRINTS; PR00007; COMPLEMENTC1Q.
DR  ProDom; PD000007; C1q_helix; 2.
DR  SMART; SM00110; C1Q; 1.
DR  PROSITE; PS01113; C1Q; 1.
KW  Extracellular matrix; Repeat; Collagen; Signal.
FT  SIGNAL      1      19      POTENTIAL.
FT  CHAIN       20     419      INNER EAR-SPECIFIC COLLAGEN.
FT  DOMAIN      20     57       NONHELICAL REGION (NC2).
FT  DOMAIN      58    274      TRIPLE-HELICAL REGION (COL1).
FT  DOMAIN      275    419      NONHELICAL REGION (NC1).
FT  DOMAIN      272    419      C1Q.
FT  CARBOHYD     37     37       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD     320    320      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ  SEQUENCE     419 AA; 43634 MW; 570CDB9675FC0F39 CRC64;

Query Match      19.1%; Score 171; DB 1; Length 419;
Best Local Similarity 31.1%; Pred. No. 1.1e-08;
Matches 59; Conservative 14; Mismatches 65; Indels 52; Gaps 8;

QY  4 DAPARGPAPP-----EP-----EPKSAFSAARTSLVGS 31
    |||||
DB  234 DTGARGPFPGRPGRGAGLKGKGLKGVGRGPKPGGESVEQIRSAFSV---GLFPS 289
    |||||

QY  32 DAGPGRHQPLAFDTEFNIGDFFDAAGVFRCLPGAYFFSF--TLGKLP-RKTL---S 85
    |||||
DB  290 RSPF-PPSLPVKFDKVFYNGEGHWDPTLNKFNVTYPGVYLLFSYHITVRNRYRALVWNG 348
    |||||

QY  86 VKLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRRGDAVWLSHDHDGYGAYSNGKTYI 145
    |||||
DB  349 VRKLRTDSLGGQIDQA-----SNLALHLTDGDQVWLETL-RDWNQYSSSEDDS 399
    |||||

QY  146 TFSGFLVYPD 155
    |||||
DB  400 TFSGFLLYPD 409
    |||||

RESULT 14
C1QC_HUMAN
ID  C1QC_HUMAN      STANDARD;      PRT;      245 AA.
AC  P02747; Q96DL2; Q96H05;
DT  21-JUL-1986 (Rel. 01, Created)
DT  28-FEB-2003 (Rel. 41, Last sequence update)
DT  15-SEP-2003 (Rel. 42, Last annotation update)
DE  Complement C1q subcomponent, C chain precursor.
GN  C1QG OR C1QC.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC  NCBI_Taxid=9606;
OX  NCBI_Taxid=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  TISSUE=Monocytes;
RX  MEDLINE=91174759; PubMed=1706597;
RA  Sellar G.C., Blake D.J., Reid K.B.M.;
RT  "Characterization and organization of the genes encoding the A-, B-
    and C-chains of human complement subcomponent C1q. The complete
    derived amino acid sequence of human C1q."
RL  Biochem. J. 274:481-490(1991).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Cerebellum;
RA  Nishi T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Suzuki Y.,
RA  Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M.,
RA  Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,

```

RA Kawakami B., Nagai K., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22386257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultly S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 29-122.
RX MEDLINE=80020137; PubMed=486087;
RA Reid K.B.M.;
RT "Complete amino acid sequences of the three collagen-like regions
RT present in subcomponent C1q of the first component of human
RL Biochem. J. 179:367-371(1979).
RN [5]
RP REVIEW OF C1Q DEFICIENCY.
RX MEDLINE=98450587; PubMed=9777412;
RA Petry F.;
RT "Molecular basis of hereditary C1q deficiency.";
RL Immunobiology 199:286-294(1998).
CC -1- FUNCTION: C1Q ASSOCIATES WITH THE PROENZYMES C1R AND C1S TO YIELD
CC C1, THE FIRST COMPONENT OF THE SERUM COMPLEMENT SYSTEM. THE
CC COLLAGEN-LIKE REGIONS OF C1Q INTERACT WITH THE CA(2+)-DEPENDENT
CC C1R(2)C1S(2) PROENZYME COMPLEX, AND EFFICIENT ACTIVATION OF C1
CC TAKES PLACE ON INTERACTION OF THE GLOBULAR HEADS OF C1Q WITH THE
CC FC REGIONS OF IGG OR IGM ANTIBODY PRESENT IN IMMUNE COMPLEXES.
CC -1- SUBUNIT: C1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q, R
CC AND S IN THE MOLAR RATION OF 1:2:2. C1Q SUBCOMPONENT IS COMPOSED
CC OF NINE SUBUNITS, SIX OF WHICH ARE DISULFIDE-LINKED DIMERS OF THE
CC A AND B CHAINS, AND THREE OF WHICH ARE DISULFIDE-LINKED DIMERS OF
CC THE C CHAIN.
CC -1- PTM: O-linked glycans consist of Glc-Gal disaccharides bound to
CC the oxygen atom of post-translationally added hydroxyl groups.
CC -1- DISEASE: Defects in C1QG are a cause of C1Q deficiency
CC [MIM:120575]. It is a rare genetic disorder which is associated
CC with recurrent infections and a high prevalence of lupus
CC erythematosus-like symptoms. It is characterized by a loss of
CC activation of the complement classical pathway.
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C1Q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AK057792; BAB71575.1; -
DR EMBL; BC009016; AAH09016.1; -
DR PIR; S14351; C1HUQC.

DR Genew; HGNC:1245; C1QG.
DR MIM; 120575; -
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003793; F:defense/immunity protein activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR00087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Complement pathway; Plasma; Hydroxylation; Glycoprotein; Collagen;
KW Repeat; Signal; Disease mutation.
KW SIGNAL 1 28
FT CHAIN 29 245
FT DOMAIN 31 112
FT DOMAIN 113 245
FT DISULFID 32 32
FT MOD_RES 36 36
FT MOD_RES 39 39
FT MOD_RES 42 42
FT MOD_RES 45 45
FT MOD_RES 54 54
FT MOD_RES 57 57
FT MOD_RES 63 63
FT MOD_RES 66 66
FT MOD_RES 71 71
FT MOD_RES 75 75
FT CARBOHYD 81 81
FT MOD_RES 84 84
FT CARBOHYD 84 84
FT MOD_RES 93 93
FT MOD_RES 96 96
FT MOD_RES 99 99
FT MOD_RES 105 105
FT MOD_RES 105 105
FT VARIANT 43 43
FT FTID=VAR 008542.
FT CONFLICT 14 14 K -> R (IN REF. 2).
FT CONFLICT 23 23 P -> A (IN REF. 1).
FT CONFLICT 57 57 K -> P (IN REF. 4).
FT CONFLICT 66 66 P -> K (IN REF. 4).
FT CONFLICT 72 72 K -> P (IN REF. 4).
FT CONFLICT 84 84 P -> K (IN REF. 4).
FT CONFLICT 87 87 N -> D (IN REF. 4).
FT CONFLICT 90 90 M -> N (IN REF. 4).
FT CONFLICT 215 215 E -> G (IN REF. 2).
SQ SEQUENCE 245 AA; 25774 MW; FA17117EB7ABFC12 CRC64;
Query Match 19.0%; Score 170; DB 1; Length 245;
Best Local Similarity 26.6%; Pred. No. 7, le-09;
Matches 46; Conservative 27; Mismatches 56; Indels 44; Gaps 6;
QY 9 GPPA-----PEP-----RSAFSAARTSLVSGDAGPGRHQP-----L 42
DB 91 GPPGMPGVPGMPGIPGEPGEGRYKQKQSVFTVTR-----QTHQPPAPNSLI 138
QY 43 AFDTEFVNIGDGFDAAGVFRCLPGAYFFSFTLGKLPKXTLSVKLMKNRDEVQAMTYDD 102
DB 139 RFNAVLTPQGDYDSTGKFTCKVPGLYYFVYHAS---HTANLCVLLYRSGVKVTFPG 194
QY 103 GASRRREMOSQSWMLALRGAVALSHDHGAYGAYSNHGKYTFSGFLVYPD 155
DB 195 HTSKTNQVNSGCVLLRLQVGEVWLAVNDY--YDMVGIGQSDSVFSGFLFPD 245
RESULT 15
CA1A CHICK STANDARD; PRT; 674 AA.
ID CA1A CHICK
AC P08125;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Collagen alpha 1(X) chain precursor.
GN COL10A1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OK NCBI_TaxID=9031;
RN [1]
RP SEQUENCE OF 48-674 FROM N.A., AND SEQUENCE OF 103-117 AND 453-466.
RX MEDLINE=86168227; PubMed=3082876;
RA Ninomiya Y., Gordon M., van der Rest M., Schmid T., Linsenmayer T.,
RA Olsen B.R.;
RT "The developmentally regulated type X collagen gene contains a long
RT open reading frame without introns."
RL J. Biol. Chem. 261:5041-5050(1986).
RN [2]
RP SEQUENCE OF 1-75 FROM N.A.
RX MEDLINE=89054019; PubMed=2461368;
RA Luvalle P., Ninomiya Y., Rosenblum N.D., Olsen B.R.;
RT "The type X collagen gene. Intron sequences split the 5'-untranslated
RT region and separate the coding regions for the non-collagenous amino-
RT terminal and triple-helical domains."
RL J. Biol. Chem. 263:18378-18385(1988).
RN [3]
RP REVISIONS TO C-TERMINUS.
RX MEDLINE=89380199; PubMed=2476437;
RA Yamaguchi N., Benya P.D., van der Rest M., Ninomiya Y.;
RT "The cloning and sequencing of alpha 1(VIII) collagen cDNAs
RT demonstrate that type VIII collagen is a short chain collagen and
RT contains triple-helical and carboxyl-terminal non-triple-helical
RT domains similar to those of type X collagen."
RL J. Biol. Chem. 264:16022-16029(1989).
CC -!- FUNCTION: TYPE X COLLAGEN IS A PRODUCT OF HYPERTROPHIC
CC CHONDROCYTES AND HAS BEEN LOCALIZED TO PRESUMPTIVE
CC MINERALIZATION ZONES OF HYALINE CARTILAGE.
CC -!- SUBUNIT: Homotrimer.
CC -!- PTM: Prolines at the third position of the tripeptide repeating
CC unit (G-X-Y) are hydroxylated in some or all of the chains.
CC -!- SIMILARITY: STRONG, TO ALPHA 1 AND 2 TYPE VIII COLLAGENS.
CC -!- SIMILARITY: Contains 1 C1Q domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M13496; AAA48736.1; ALT_SEQ.
DR EMBL; J04194; AAA48634.1; -.
DR PIR; S23297; S23297.
DR InterPro; IPR001073; C1Q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1Q; 1.
DR Pfam; PF01391; Collagen; 8.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR ProDom; PD000007; C1q_helix; 1.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Cartilage; Collagen; Signal.
FT SIGNAL 1 18
FT CHAIN 19 674 COLLAGEN ALPHA 1(X) CHAIN.
FT DOMAIN 19 52 NONHELICAL REGION (NC2).
FT DOMAIN 53 512 TRIPLE-HELICAL REGION.
FT DOMAIN 513 674 NONHELICAL REGION (NC1).
FT DOMAIN 539 674 C1Q.
FT MOD_RES 453 453 HYDROXYLATION.
FT MOD_RES 456 456 HYDROXYLATION.
SQ SEQUENCE 674 AA; 66434 MW; EAB48B1EF174B145 CRC64;

Query Match 18.6%; Score 166.5; DB 1; Length 674;
Best Local Similarity 28.9%; Pred. No. 5e-08;
Matches 50; Conservative 19; Mismatches 71; Indels 33; Gaps 5;
QY 9 GPPAPPEP-----RSAFSAARTSLVG-----SDAGPGR 38
DB 501 GPPGPPGPPGGSTIPEGYKGESRELSGMSFMKAGANQALTGMPVSAFYILSKAYPGAT 560
QY 39 HQPLAFDTEFVNIIGDDEDAAGVFCRLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAM 98
DB 561 -VPIKFDKILYNRQOHYDPRGTGIFTCRIPGLYFSYHV-HAKGTNVWVALYKNGSPVM-Y 617
QY 99 IYDDGASRRREMOSQSVMLALRRGDAVWLLSHDGYGAYSNHGKYITFSGFL 151
DB 618 TYDEYOKGYLDQASGSAVIDLMENDQVWLQEPNSESNGLYSSEYVHSSFSGFL 670

Search completed: January 12, 2004, 08:12:59
Job time : 7.10778 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:10 ; Search time 23.4132 Seconds
(without alignments)
1873.686 Million cell updates/sec

Title: US-10-085-167-2_COPY_160_358
Perfect score: 894
Sequence: 1 ADADAPARGPPAPPEPRSAF.....LYPPDLAPAPPGIGASELL 170

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 segs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 894 | 100.0 | 329 | 4 Q8IV25 | Q8IV25 homo sapien |
| 2 | 807.5 | 90.3 | 205 | 11 Q9DOW2 | Q9DOW2 mus musculu |
| 3 | 807.5 | 90.3 | 326 | 11 Q8R066 | Q8R066 mus musculu |
| 4 | 801.5 | 89.7 | 205 | 11 Q9DCB6 | Q9DCB6 mus musculu |
| 5 | 230 | 25.7 | 158 | 4 Q9H667 | Q9H667 homo sapien |
| 6 | 230 | 25.7 | 1077 | 4 Q8TE71 | Q8TE71 homo sapien |
| 7 | 226 | 25.3 | 158 | 11 Q8K110 | Q8K110 mus musculu |
| 8 | 220.5 | 24.7 | 182 | 11 Q8R1P2 | Q8R1P2 mus musculu |
| 9 | 220.5 | 24.7 | 281 | 11 Q9QXP7 | Q9QXP7 mus musculu |
| 10 | 208.5 | 23.3 | 243 | 4 Q8N6P2 | Q8N6P2 homo sapien |
| 11 | 197.5 | 22.1 | 243 | 11 Q8R002 | Q8R002 mus musculu |
| 12 | 194.5 | 21.8 | 243 | 11 Q8K479 | Q8K479 mus musculu |
| 13 | 189 | 21.1 | 247 | 11 Q8BRW2 | Q8BRW2 mus musculu |
| 14 | 187.5 | 21.0 | 294 | 11 Q9D8U4 | Q9D8U4 mus musculu |
| 15 | 185 | 20.7 | 240 | 6 Q95MQ4 | Q95mq4 bos taurus |
| 16 | 180 | 20.1 | 244 | 11 Q8K3R4 | Q8K3r4 rattus norv |

| | | | | | |
|----|-------|------|------|-----------|--------------------|
| 17 | 179.5 | 20.1 | 120 | 11 Q8R1Z2 | Q8r1z2 mus musculu |
| 18 | 179.5 | 20.1 | 264 | 11 Q8BKRO | Q8bkro mus musculu |
| 19 | 177 | 19.8 | 287 | 11 Q8CFR0 | Q8cfro mus musculu |
| 20 | 176 | 19.7 | 243 | 6 Q95JD7 | Q95jd7 macaca mula |
| 21 | 166 | 18.6 | 675 | 6 Q9N178 | Q9n178 sus scrofa |
| 22 | 163 | 18.2 | 224 | 4 Q8IUK8 | Q8luk8 homo sapien |
| 23 | 163 | 18.2 | 224 | 11 Q8BGU2 | Q8bgu2 mus musculu |
| 24 | 161 | 18.0 | 173 | 6 Q62789 | Q62789 sus scrofa |
| 25 | 161 | 18.0 | 246 | 11 Q9ES30 | Q9es30 mus musculu |
| 26 | 160.5 | 18.0 | 295 | 11 Q9Z1K4 | Q9z1k4 rattus norv |
| 27 | 157.5 | 17.6 | 744 | 11 Q921S8 | Q921s8 mus musculu |
| 28 | 157.5 | 17.6 | 744 | 11 Q8BGL6 | Q8bgl6 mus musculu |
| 29 | 152 | 17.0 | 197 | 11 Q9JHG0 | Q9jhg0 mus musculu |
| 30 | 152 | 17.0 | 312 | 11 Q8CHX9 | Q8chx9 mus musculu |
| 31 | 151 | 16.9 | 194 | 6 Q95J95 | Q95j95 canis famil |
| 32 | 148 | 16.6 | 196 | 11 Q920N0 | Q920n0 tarias sibi |
| 33 | 147.5 | 16.5 | 333 | 4 Q8IUN4 | Q8iun4 homo sapien |
| 34 | 143 | 16.0 | 289 | 11 Q8BVD7 | Q8bvd7 mus musculu |
| 35 | 142 | 15.9 | 705 | 4 Q8TEJ5 | Q8tej5 homo sapien |
| 36 | 134 | 15.0 | 245 | 11 Q9DCM6 | Q9dcm6 mus musculu |
| 37 | 134 | 15.0 | 1017 | 11 Q99K41 | Q99k41 mus musculu |
| 38 | 133.5 | 14.9 | 195 | 11 Q8BZS3 | Q8bzs3 mus musculu |
| 39 | 130 | 14.5 | 164 | 11 Q64144 | Q64144 rattus sp. |
| 40 | 129 | 14.4 | 347 | 4 Q96IH6 | Q96ih6 homo sapien |
| 41 | 129 | 14.4 | 583 | 4 Q96G58 | Q96g58 homo sapien |
| 42 | 129 | 14.4 | 992 | 4 Q9UG76 | Q9ug76 homo sapien |
| 43 | 129 | 14.4 | 1016 | 4 Q9Y6C2 | Q9y6c2 homo sapien |
| 44 | 127.5 | 14.3 | 198 | 11 Q8BMF0 | Q8bmf0 mus musculu |
| 45 | 126.5 | 14.1 | 198 | 11 Q8BME9 | Q8bme9 mus musculu |

ALIGNMENTS

| RESULT 1 | ID | Q8IV25 | PRELIMINARY; | PRT; | 329 AA. |
|----------|---|---|--------------|------------------|---------|
| AC | Q8IV25 | Q8IV25 | | | |
| DT | 01-MAR-2003 | (TREMBLrel. 23, Created) | | | |
| DT | 01-MAR-2003 | (TREMBLrel. 23, Last sequence update) | | | |
| DT | 01-MAR-2003 | (TREMBLrel. 23, Last annotation update) | | | |
| DE | Similar to Clq and tumor necrosis factor related protein 4. | | | | |
| OS | Homo sapiens (Human). | | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. | | | | |
| OX | NCBI_TaxID=9606; | | | | |
| RN | [1] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RC | TISSUE=Brain; | | | | |
| RA | Strausberg R.; | | | | |
| RL | Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases. | | | | |
| DR | EMBL; BC035628; AAH35628.1; - | | | | |
| SQ | SEQUENCE | 329 AA; | 35256 MW; | 16064DA8182A6732 | CRG64; |

Query Match 100.0%; Score 894; DB 4; Length 329;
Best Local Similarity 100.0%; Pred. No. 2.9e-82;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----------|-----|---|-----|
| QY | 1 | ADADAPARGPPAPPEPRSAFSAARTRSLVSGDAGPGRHOPLAFDTEFVNIGDFDAAG | 60 |
| DB | 160 | ADADAPARGPPAPPEPRSAFSAARTRSLVSGDAGPGRHOPLAFDTEFVNIGDFDAAG | 219 |
| QY | 61 | VFRGRUGAYFFSFTLGKLPKTLVKLMKNRDEVOAMIVDGAARRRMOSSQSVMLALR | 120 |
| DB | 220 | VFRGRUGAYFFSFTLGKLPKTLVKLMKNRDEVOAMIVDGAARRRMOSSQSVMLALR | 279 |
| QY | 121 | RGDAVWLISHDHGAYSNHGYITFSGLVYPDLAPAPPGIGASELL | 170 |
| DB | 280 | RGDAVWLISHDHGAYSNHGYITFSGLVYPDLAPAPPGIGASELL | 329 |
| RESULT 2 | | | |
| Q9DOW2 | | | |

ID Q9D0W2 PRELIMINARY; PRT; 205 AA.
AC Q9D0W2;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE 0710001E10R1k protein.
GN 0710001E10R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staabl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-Oka K., Wang K.H., Weltz C., Wittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK004340; BAB23268.1; "
DR MGD; MGI:1914695; 0710001E10R1k.
DR InterPro; IPR001073; Clq.
DR Pfam; PF00386; Clq; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; Clq; 1.
DR PROSITE; PS01113; Clq; 1.
SQ SEQUENCE 205 AA; 22190 MW; B9B237793C54786D CRC64;

Query Match 90.3%; Score 807.5; DB 11; Length 205;
Best Local Similarity 92.9%; Pred. No. 9.2e-74;
Matches 157; Conservative 1; Mismatches 10; Indels 1; Gaps 1;
QY 1 ADADAPARGPPAPPEPRSAFSAARTSLVSGDAGPPRHQPLAFDTEFVNIGDFDAAAG 60
DB 38 ADADAPARG-PAAPEPRSAFSAARTSLVSGSDAAGPRHRPLAFDTELVNIGDFDAAAG 96
QY 61 VFCRLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALR 120
DB 97 VFCRLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLPLR 156
QY 121 RGDVWLLSHDHGCGAYSNHGKTYTFSGFLVYPPDLAPAPPGIGASEL 169
DB 157 RGDVWLLSHDHGCGAYSNHGKTYTFSGFLVYPPDLAAGPALKPPEL 205

RESULT 3
Q8R066 PRELIMINARY; PRT; 326 AA.
ID Q8R066;
AC Q8R066;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Similar to Clq and tumor necrosis factor related protein 4.
GN 0710001E10R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027315; AAH27315.1; "
DR MGD; MGI:1914695; 0710001E10R1k.
DR InterPro; IPR001073; Clq.
DR Pfam; PF00386; Clq; 2.
DR SMART; SM00110; Clq; 2.
DR PROSITE; PS01113; Clq; 2.
SQ SEQUENCE 326 AA; 35057 MW; 72339172B7B1051A CRC64;

Query Match 90.3%; Score 807.5; DB 11; Length 326;
Best Local Similarity 92.9%; Pred. No. 1.7e-73;
Matches 157; Conservative 1; Mismatches 10; Indels 1; Gaps 1;
QY 1 ADADAPARGPPAPPEPRSAFSAARTSLVSGDAGPPRHQPLAFDTEFVNIGDFDAAAG 60
DB 159 ADADAPARG-PAAPEPRSAFSAARTSLVSGSDAAGPRHRPLAFDTELVNIGDFDAAAG 217
QY 61 VFCRLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALR 120
DB 218 VFCRLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLPLR 277
QY 121 RGDVWLLSHDHGCGAYSNHGKTYTFSGFLVYPPDLAPAPPGIGASEL 169
DB 278 RGDVWLLSHDHGCGAYSNHGKTYTFSGFLVYPPDLAAGPALKPPEL 326

RESULT 4
Q9DCB6 PRELIMINARY; PRT; 205 AA.
ID Q9DCB6;
AC Q9DCB6;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE 0710001E10R1k protein.
GN 0710001E10R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staabl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-Oka K., Wang K.H., Weltz C., Wittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK002948; BAB22473.1; "
DR MGD; MGI:1914695; 0710001E10R1k.
DR InterPro; IPR001073; Clq.
DR Pfam; PF00386; Clq; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; Clq; 1.

```
DR PROSITE; PS01113; C1Q; 1.
SQ SEQUENCE 205 AA; 22215 MW; 56AD37793C437300 CRC64;

Query Match
Best Local Similarity 89.7%; Score 801.5; DB 11; Length 205;
Matches 156; Conservative 1; Mismatches 11; Indels 1; Gaps 1;

QY 1 ADADAPARGPPAPPEPRSAFSAARTSLVSGDAGPGPRHQPLAFDTEFVNIGDPAAG 60
Db 38 ADADAPARG-PAAPEPRSAFSAARTSLVSGDAAPGPRHPLAFDTELVNIGDPAAG 96
QY 61 VERCRLPGAYFFSFTLGKLPKRLSVKLMKNRDEVQAMTYDDGASRRREMOSQSVMLALR 120
Db 97 VERCRLPGAYFFSFTLGKLPKRLSVKLMKNRDEVQAMTYDDGASRRREMOSQSVRLPLR 156
QY 121 RGDVWVLLSHDHDGYGAYSNHGKTYTFSGFLVYPPDLAPAAPGLGASEL 169
Db 157 RGDVWVLLSHDHDGYGAYSNHGKTYTFSGFLVYPPDLAPAAPGLKPEL 205
```

RESULT 5

```
Q9H667 PRELIMINARY; PRT; 158 AA.
ID Q9H667;
AC Q9H667;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Hypothetical protein FLJ22569.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK026222; BAB15398.1; -
DR EMBL; BC007520; AAH07520.1; -
DR InterPro; IPR001073; C1q.
DR Pfam; PF00386; C1q; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Hypothetical protein.
SQ SEQUENCE 158 AA; 17625 MW; 47DB10EDD6DC9760 CRC64;
```

```
Query Match
Best Local Similarity 25.7%; Score 230; DB 4; Length 158;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

QY 4 DAPARGP-----PAPPEPRSAFSAARTSLVSGDAGPGPRHQPLAFDTEFVNIG 53
Db 5 DVPVTNPAATILPVHYPLPQOMRVAFAAART-----SNLAPGTLDPPIVFDLLNLNGE 59
QY 54 DFDAAGVFERCRLPGAYFFSFTLGKLP-RKTLGVKLMKNRDEVQAMTYDDGASRRREMOS 112
Db 60 TFDLQLGRFNCVPNGTYVFIHMLKLAIVNPLVYVLMKNNEVLVSAYANDGAP-DHETAS 118
QY 113 QSVMLALRRGDVWVLLSHDHDGYGA-YSNHGKTYTFSGFLVYPPD 155
Db 119 NHAIIQLFGDQIWLRLH----RGAIGSSWKXYSTFSGYLLYQD 158
```

RESULT 6

```
Q8TE71 PRELIMINARY; PRT; 1077 AA.
ID Q8TE71
AC Q8TE71;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE EGG1.
GN EGG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Aerbadjnal W., Miller J.L.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY074490; AAL71549.1; -
DR InterPro; IPR001073; C1q.
DR Pfam; PF00386; C1q; 1.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
SQ SEQUENCE 1077 AA; 120974 MW; 2B88BF3C47D032D6 CRC64;
```

```
Query Match
Best Local Similarity 25.7%; Score 230; DB 4; Length 1077;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;
```

```
QY 4 DAPARGP-----PAPPEPRSAFSAARTSLVSGDAGPGPRHQPLAFDTEFVNIG 53
Db 924 DVPVTNPAATILPVHYPLPQOMRVAFAAART-----SNLAPGTLDPPIVFDLLNLNGE 978
QY 54 DFDAAGVFERCRLPGAYFFSFTLGKLP-RKTLGVKLMKNRDEVQAMTYDDGASRRREMOS 112
Db 979 TFDLQLGRFNCVPNGTYVFIHMLKLAIVNPLVYVLMKNNEVLVSAYANDGAP-DHETAS 1037
QY 113 QSVMLALRRGDVWVLLSHDHDGYGA-YSNHGKTYTFSGFLVYPPD 155
Db 1038 NHAIIQLFGDQIWLRLH----RGAIGSSWKXYSTFSGYLLYQD 1077
```

RESULT 7

```
Q8K110 PRELIMINARY; PRT; 158 AA.
ID Q8K110
AC Q8K110;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Similar to hypothetical protein FLJ22569.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027523; AAH27523.1; -
DR InterPro; IPR001073; C1q.
DR Pfam; PF00386; C1q; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Hypothetical protein.
SQ SEQUENCE 158 AA; 17533 MW; 86E9321C99225FCB CRC64;
```

```
Query Match
Best Local Similarity 25.3%; Score 226; DB 11; Length 158;
Matches 62; Conservative 17; Mismatches 63; Indels 22; Gaps 6;
```

```
QY 4 DAPARGP-----PAPPEPRSAFSAARTSLVSGDAGPGPRHQPLAFDTEFVNIG 53
Db 5 DVPVTSPAAILPVHYPLPQOMRVAFAAART-----SNLAPGTLDPPIVFDLLNLNGE 59
```


[illegible]

| ID | Q8R002 | PRELIMINARY; | PRT; | 243 AA. |
|----|--|------------------|--------------------|-------------------------------------|
| AC | Q8R002; | | | |
| DT | 01-JUN-2002 (TREMBLrel. 21, Created) | | | |
| DT | 01-JUN-2002 (TREMBLrel. 21, Last sequence update) | | | |
| DT | 01-MAR-2003 (TREMBLrel. 23, Last annotation update) | | | |
| DE | Similar to DKFZP586B0621 protein (Hypothetical 25.4 kDa protein). | | | |
| GN | C1QTNF5. | | | |
| OS | Mus musculus (Mouse). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | |
| OX | NCBI_TaxID=10090; | | | |
| RN | (11) | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | Strausberg R.; | | | |
| RL | Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases. | | | |
| RP | (2) | | | |
| RA | SEQUENCE FROM N.A. | | | |
| RA | Strausberg R.; | | | |
| RL | Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases. | | | |
| DR | EMBL; BC023068; AAH23068.1; - | | | |
| DR | EMBL; BC025174; AAH25174.1; - | | | |
| DR | MGD; MGI:2385958; C1qtnf5. | | | |
| DR | InterPro; IPR001073; C1q. | | | |
| DR | InterPro; IPR000087; Collagen. | | | |
| DR | Pfam; PF00386; C1q; 1. | | | |
| DR | Pfam; PF01391; Collagen; 1. | | | |
| DR | SMART; SM00110; C1Q; 1. | | | |
| KW | Hypothetical protein. | | | |
| SQ | SEQUENCE 243 AA; 25420 MW; 498129CD051DB97B CRC64; | | | |
| | Query Match | 22.1%; | Score 197.5; | DB 11; Length 243; |
| | Best local Similarity | 33.3%; | Pred. No. 7.1e-12; | |
| | Matches 55; | Conservative 16; | Mismatches 71; | Indels 23; Gaps 3; |
| QY | 8 RGPAP----- | -----PEPSAFAA | TRSLVGS | DAGPGRHQPLAFDTEFVN 50 |
| DB | 77 RGEPRGPRGEAGPMGAIGPAGEC | SVPPRS | SAFS | AKRSSESrv-----PPPADTPLPFDRLVLLN 131 |
| QY | 51 IGDPDFDAAGVFRCLPGAYFFS | FTLGKLP | RKTL | SVKLMKNRDEVQAMTYDDGASRRREM 110 |
| DB | 132 EQGHFDPTTGKFTCGVPGVYYFA | -VHATV | YRASLQ | FDLYKNGOSIASFQYFGGMPKPAS 190 |
| QY | 111 QSQSVMLALRRGDAVWL | LSHDH | DGYG | AYSNHGYITFSGFLVYPD 155 |
| DB | 191 LSGAMVRLPEPDQVWVQVGV | VDYIGI | YASIK | TDSTFSGFLVYSD 235 |

RESULT 12
Q8K479

```

ID Q8K479 PRELIMINARY; PRT; 243 AA.
AC Q8K479;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Complement-clq tumor necrosis factor-related protein.
GN C1QTNF5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
XP [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22135657; PubMed=12140190;
RA Kameya S., Hawes N.L., Chang B., Heckenlively J.R., Naggert J.K.,
RA Nishina P.M.;
RT "Mftrp, a gene encoding a frizzled related protein, is mutated in the
RT mouse retinal degeneration 6.";
RL Hum. Mol. Genet. 11:1879-1886(2002).
DR EMBL; AF469650; AA089217.1; -.
DR MGD; MGI:2385958; C1qtnf5.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 2.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
SQ SEQUENCE 243 AA; 25436 MW; 9F4D5804349791D9 CRC64;

```

```

Query Match      21.8%; Score 194.5; DB 11; Length 243;
Best Local Similarity 32.7%; Pred. No. 1.4e-11;
Matches 54; Conservative 17; Mismatches 71; Indels 23; Gaps 3;

QY      8 RGPAP-----PEPSAFSAARTRSLVGSADAGPPRHQPLAFDTEFVN 50
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      77 RGEPPRGEAGPMGAIGPAGECSVPPRSASFSAKRSERV-----PPADTPLPFDRVLIN 131
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      51 IGDFDAAGVFCRLPGAYFFSFTLGKLPKTLVSKLTKNRDEYQAMLYDDGASRRREM 110
      | : | : | : | : | : | : | : | : | : | : | : | : | :
Db      132 EQGHYDPTTGKFTCQVPGVYFA-VHATVYRASLQFDLVKNGQSLASFQYFGGWPKPAS 190
      | : | : | : | : | : | : | : | : | : | : | : | : | :

QY      111 QSQSVMLALRRGDAVWMLSHDHDGYGAYSNHGKIYTFSGFLVYPD 155
      | : | : | : | : | : | : | : | : | : | : | : | : | :
Db      191 LSGAMVRLPEPDQVWVQVGVDYIGIYASIKTDSTFSGFLVYSD 235
      | : | : | : | : | : | : | : | : | : | : | : | : | :

```

```

RESULT 13
Q8BRW2
ID Q8BRW2 PRELIMINARY; PRT; 247 AA.
AC Q8BRW2;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Adipocyte complement related protein of 30 kDa.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=22354683; Pubmed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK041214; BAC30866.1; -.
SQ SEQUENCE 247 AA; 26751 MW; 0D3FA64C789CAEF3 CRC64;

```

| | | | | |
|-----------------------|--------|--------------------|--------|-------------|
| Query Match | 21.1%; | Score 189; | DB 11; | Length 247; |
| Best Local Similarity | 33.5%; | Pred. No. 5.3e-11; | | |

RESULT 12
Q8K479

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:09 ; Search time 25.6188 Seconds
(without alignments)
1053.272 Million cell updates/sec

Title: US-10-085-167-2_COPY_160_358
Perfect score: 894
Sequence: 1 ADADAPARGPPAPPEPRSAF.....LVYPDLAPAAPGLGASELL 170

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*
1: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgcdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|--------------------|
| 1 | 894 | 100.0 | 329 | 22 | AAB61606 | Human ZACRP4. Hom |
| 2 | 894 | 100.0 | 329 | 23 | ABG79643 | Human novel secret |
| 3 | 871 | 97.4 | 329 | 22 | AAB61424 | Monkey MANGO 245 p |
| 4 | 743 | 83.1 | 192 | 22 | AAB61479 | Human MANGO 245 cy |
| 5 | 743 | 83.1 | 334 | 22 | AAB61466 | Human MANGO 245 ma |
| 6 | 743 | 83.1 | 348 | 22 | AAB61423 | Human MANGO 245 pr |
| 7 | 730 | 81.7 | 299 | 23 | ABG70385 | Adipocyte compleme |
| 8 | 696 | 77.9 | 221 | 23 | ABG70383 | Adipocyte compleme |
| 9 | 683 | 76.4 | 134 | 22 | AAB61473 | Monkey MANGO 245 C |

| | | | | | | |
|----|-------|------|------|----|----------|--------------------|
| 10 | 666.5 | 74.6 | 284 | 23 | ABG70384 | Adipocyte compleme |
| 11 | 664 | 74.3 | 127 | 22 | ABB11578 | Human secreted pro |
| 12 | 609 | 68.1 | 117 | 22 | AAB61470 | Human MANGO 245 C1 |
| 13 | 482 | 53.9 | 130 | 22 | AAB61468 | Mature monkey MANG |
| 14 | 432 | 48.3 | 133 | 23 | ABG70386 | Adipocyte compleme |
| 15 | 357.5 | 40.0 | 199 | 22 | AAB61488 | Murine MANGO 245 p |
| 16 | 321 | 35.9 | 126 | 22 | AAB61472 | Monkey MANGO 245 C |
| 17 | 320 | 35.8 | 125 | 22 | AAB61477 | Human MANGO 245 ex |
| 18 | 320 | 35.8 | 126 | 22 | AAB61469 | Human MANGO 245 C1 |
| 19 | 318 | 35.6 | 126 | 22 | AAB61489 | Murine MANGO 245 C |
| 20 | 305.5 | 34.2 | 123 | 21 | AAB42189 | Human ORFX ORF1953 |
| 21 | 237.5 | 26.6 | 252 | 22 | AAB49590 | Human zacrpf. Hom |
| 22 | 237.5 | 26.6 | 252 | 23 | AAO19412 | Human molecule for |
| 23 | 237.5 | 26.6 | 252 | 23 | AAU83100 | Novel secreted pro |
| 24 | 237.5 | 26.6 | 252 | 23 | ABB53292 | Human polypeptide |
| 25 | 230 | 25.7 | 202 | 22 | AAU99927 | Human polypeptide |
| 26 | 230 | 25.7 | 709 | 23 | ABB08645 | Human pancreatic c |
| 27 | 230 | 25.7 | 710 | 23 | AAO15420 | Human ganset metab |
| 28 | 230 | 25.7 | 746 | 20 | AAU29512 | Human lung tumour |
| 29 | 230 | 25.7 | 746 | 21 | AAU44461 | Human lung tumour |
| 30 | 230 | 25.7 | 746 | 22 | AAU13802 | Human lung tumour |
| 31 | 230 | 25.7 | 800 | 22 | AAU78334 | Human protein SEO |
| 32 | 230 | 25.7 | 908 | 22 | AAE09839 | Novel human protei |
| 33 | 230 | 25.7 | 909 | 22 | AAE09845 | Novel human protei |
| 34 | 230 | 25.7 | 957 | 22 | AAE09841 | Novel human protei |
| 35 | 230 | 25.7 | 958 | 22 | AAE09847 | Novel human protei |
| 36 | 230 | 25.7 | 992 | 22 | AAE09843 | Novel human protei |
| 37 | 230 | 25.7 | 993 | 22 | AAE09849 | Novel human protei |
| 38 | 230 | 25.7 | 1042 | 22 | AAE09838 | Novel human protei |
| 39 | 230 | 25.7 | 1043 | 22 | AAE09844 | Novel human protei |
| 40 | 230 | 25.7 | 1091 | 22 | AAE09840 | Novel human protei |
| 41 | 230 | 25.7 | 1092 | 22 | AAE09846 | Novel human protei |
| 42 | 230 | 25.7 | 1126 | 22 | AAE09842 | Novel human protei |
| 43 | 230 | 25.7 | 1127 | 22 | AAE09848 | Novel human protei |
| 44 | 220.5 | 24.7 | 281 | 20 | AAU98014 | Mouse adipocyte-sp |
| 45 | 220.5 | 24.7 | 281 | 21 | AAU08428 | Amino acid sequenc |

ALIGNMENTS

| | | | | | |
|----------|--|----------------------------|--|--|--|
| RESULT 1 | | | | | |
| AAB61606 | | | | | |
| ID | AAB61606 | standard; Protein; 329 AA. | | | |
| XX | XX | | | | |
| AC | AAB61606; | | | | |
| XX | XX | | | | |
| DT | 05-APR-2001 | (first entry) | | | |
| XX | XX | | | | |
| DE | Human ZACRP4. | | | | |
| XX | XX | | | | |
| KW | Human; zacrpf; complement factor C1q domain; chromosome 11q11; | | | | |
| KW | energy balance; cellular metabolic reaction; autocrine factor; | | | | |
| KW | development; cell proliferation; differentiation; cell survival. | | | | |
| XX | XX | | | | |
| OS | Homo sapiens. | | | | |
| XX | XX | | | | |
| FH | Key | Location/Qualifiers | | | |
| FT | Peptide | 1..16 | | | |
| FT | FT | /label= Signal_peptide | | | |
| FT | Domain | 17..159 | | | |
| FT | FT | /label= C1q_domain_#1 | | | |
| FT | Domain | 160..328 | | | |
| FT | FT | /label= C1q_domain_#2 | | | |
| XX | XX | | | | |
| PN | WO200102565-A2. | | | | |
| XX | XX | | | | |
| PD | 11-JAN-2001. | | | | |
| XX | XX | | | | |
| PF | 28-JUN-2000; 2000WO-US17692. | | | | |
| XX | XX | | | | |
| PR | 01-JUL-1999; 99US-0346502. | | | | |
| XX | XX | | | | |

PA (ZYMO) ZYMOGENETICS INC.
XX
PI Holloway JL, Lok S;
XX
DR WPI; 2001-138140/14.
DR N-PSDB; AAF28672.
XX
PT Novel secreted protein ZACRP4 polypeptides having tandem C1q globular
PT domains, useful for studying cell-cell communication and regulation of
PT cellular processes -
XX
PS Claim 1; Page 77-78; 82pp; English.
XX
CC The present sequence is human ZACRP4 protein. ZACRP4 protein has two
CC complement factor C1q domains. The ZACRP4 gene is located on human
CC chromosome 1q11. The ZACRP4 coding sequence and protein have a number of
CC uses described in the specification, including, modulation of energy
CC balance and cellular metabolic reactions in mammals. In addition, ZACRP4
CC protein is useful as an autocrine factor, particularly during
CC development, in mediating the processes of an organism, in regulating
CC cellular processes such as cell proliferation and/or differentiation,
CC cell survival and energy balance.
XX
SQ Sequence 329 AA;

Query Match 100.0%; Score 894; DB 22; Length 329;
Best Local Similarity 100.0%; Pred. No. 3.9e-89;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADADAPARGPPAPPEPRSAFSAARTSLVSGDAGGPRHQPLAFDTEFVNIGDFDAAG 60
DB 160 ADADAPARGPPAPPEPRSAFSAARTSLVSGDAGGPRHQPLAFDTEFVNIGDFDAAG 219
QY 61 VFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALR 120
DB 220 VFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALR 279
QY 121 RGDVWLLSHDHDGYGAYSNHGKYYITFSGLVYPDLAPAPPGLGASELL 170
DB 280 RGDVWLLSHDHDGYGAYSNHGKYYITFSGLVYPDLAPAPPGLGASELL 329

RESULT 2
ABG79643
ID ABG79643 standard; Protein; 329 AA.
XX
AC ABG79643;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human novel secreted protein SECP19, Incyte ID No. 931619CD1.
XX
KW Human; SECP; secreted protein; cell proliferative disorder;
KW actinic keratosis; arteriosclerosis; bursitis; hepatitis; cancer;
KW autoimmune disorder; inflammatory disorder; AIDS; asthma; allergy;
KW acquired immunodeficiency syndrome; anaemia; atopic dermatitis;
KW cardiovascular disorder; congestive heart failure; vascular tumour;
KW ischaemic heart disease; myocardial infarction; epilepsy; stroke;
KW hypertensive heart disease; neurological disorder; cerebral neoplasm;
KW Alzheimer's disease; developmental disorder; renal tubular acidosis;
KW Cushing's syndrome; Duchenne muscular dystrophy; hypothyroidism;
KW Becker muscular dystrophy.
XX
OS Homo sapiens.
XX
PN WO200262841-A2.
XX
PD 15-AUG-2002.
XX
PF 28-JAN-2002; 2002WO-US02616.
XX
PR 02-FEB-2001; 2001US-266195P.
PR 08-FEB-2001; 2001US-267924P.
XX

PR 09-FEB-2001; 2001US-267816P.
PR 09-FEB-2001; 2001US-268112P.
PR 26-FEB-2001; 2001US-271639P.
PR 07-SEP-2001; 2001US-317818P.
PR 21-DEC-2001; 2001US-343553P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Tang TY, Yue H, Gandhi AR, Yao MG, Warren BA, Ding L, Duggan BM;
PI Xu Y, Yang J, Thangavelu K, Lal PG, Honchell CD, Walia NK, Lee S;
PI Lee EA, Richardson TW, Baughn MR, Elliott VS;
XX
DR WPI; 2002-657522/70.
DR N-PSDB; ABS64954.
XX
PT New human secreted proteins and nucleic acids useful in diagnosing,
PT treating and preventing cell proliferative, autoimmune/inflammatory,
PT cardiovascular, neurological, and developmental disorders -
XX
PS Claim 1; Page 140; 158pp; English.
XX
CC The invention relates to twenty four human secreted proteins
CC (SECP1-24), proteins 90% identical to them and active fragments of them.
CC Also included are nucleic acids encoding the SECP proteins, a recombinant
CC polynucleotide comprising a promoter sequence operably linked to the
CC nucleic acid, a cell transformed with the recombinant polynucleotide, an
CC a transgenic organism comprising the recombinant polynucleotide, an
CC anti-SECP antibody, and screening for ant/agonists and modulators of
CC SECP function or expression. The SECP proteins and nucleic acids are
CC useful in the diagnosis, treatment and prevention of cell proliferative
CC (e.g. actinic keratosis, arteriosclerosis, bursitis, hepatitis or
CC cancer), autoimmune/inflammatory (e.g. AIDS (acquired immunodeficiency
CC syndrome), asthma, anaemia, allergies or atopic dermatitis),
CC cardiovascular (e.g. congestive heart failure, ischaemic heart disease,
CC myocardial infarction, hypertensive heart disease, or vascular tumours),
CC neurological (e.g. epilepsy, stroke, cerebral neoplasms, or Alzheimer's
CC disease), and developmental (e.g. renal tubular acidosis, Cushing's
CC syndrome, Duchenne and Becker muscular dystrophy, or hypothyroidism)
CC disorders. Many other diseases and disorders are listed in the
CC specification. These may also be used in assessing the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of the secreted proteins. The present sequence represents a
CC SECP protein of the invention.
XX
SQ Sequence 329 AA;

Query Match 100.0%; Score 894; DB 23; Length 329;
Best Local Similarity 100.0%; Pred. No. 3.9e-89;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADADAPARGPPAPPEPRSAFSAARTSLVSGDAGGPRHQPLAFDTEFVNIGDFDAAG 60
DB 160 ADADAPARGPPAPPEPRSAFSAARTSLVSGDAGGPRHQPLAFDTEFVNIGDFDAAG 219
QY 61 VFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALR 120
DB 220 VFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALR 279
QY 121 RGDVWLLSHDHDGYGAYSNHGKYYITFSGLVYPDLAPAPPGLGASELL 170
DB 280 RGDVWLLSHDHDGYGAYSNHGKYYITFSGLVYPDLAPAPPGLGASELL 329

RESULT 3
AAB61424
ID AAB61424 standard; protein; 329 AA.
XX
AC AAB61424;
XX
DT 04-APR-2001 (first entry)
XX
DE Monkey MANGO 245 protein.
XX

KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Catarhini sp..
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Claim 1; Fig 25; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 329 AA;

Query Match 97.4%; Score 871; DB 22; Length 329;
Best Local Similarity 97.6%; Pred. No. 1.3e-86;
Matches 166; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ADADAPARGPAPPEPRSAFSAARTSLVGS DAGPGPRHQPLAFDTEFVNIGDFDAAG 60
Db 160 ADADAPARGPAPPEPRSAFSAARTSLVGS DAGGPRHRPLAFDTELVNIGDFDAAG 219

QY 61 VFCRLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALR 120
Db 220 VFCRLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALR 279

QY 121 RGDVWVLLSHDHDGYGAYSNHGKYITFSGFLVYPDLAPAPPGIGASELL 170
Db 280 RGDVWVLLSHDHDGYGAYSNHGKYITFSGFLVYPDLAPAPPGIGASELL 329

RESULT 4
AAB61479
ID AAB61479 standard; protein; 192 AA.
XX
AC AAB61479;
XX
DT 04-APR-2001 (first entry)
XX
DE Human MANGO 245 cytoplasmic domain.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Homo sapiens.
XX
PN WO200100672-A1.
XX

PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Disclosure; Page 247; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 192 AA;

Query Match 83.1%; Score 743; DB 22; Length 192;
Best Local Similarity 100.0%; Pred. No. 6.2e-73;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADADAPARGPAPPEPRSAFSAARTSLVGS DAGPGPRHQPLAFDTEFVNIGDFDAAG 60
Db 4 ADADAPARGPAPPEPRSAFSAARTSLVGS DAGPGPRHQPLAFDTEFVNIGDFDAAG 63

QY 61 VFCRLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALR 120
Db 64 VFCRLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALR 123

QY 121 RGDVWVLLSHDHDGYGAYSNH 141
Db 124 RGDVWVLLSHDHDGYGAYSNH 144

RESULT 5
AAB61466
ID AAB61466 standard; protein; 334 AA.
XX
AC AAB61466;
XX
DT 04-APR-2001 (first entry)
XX
DE Human MANGO 245 mature protein.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Homo sapiens.
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX

PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Disclosure; Page 242; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 334 AA;
Query Match 83.1%; Score 743; DB 22; Length 334;
Best Local Similarity 100.0%; Pred. No. 1.2e-72;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ADADAPARGPPAPPEPRSAFSAARTSLVGS DAGPGRHQPLAFDTEFVNIGD FDDAAAG 60
Db 146 ADADAPARGPPAPPEPRSAFSAARTSLVGS DAGPGRHQPLAFDTEFVNIGD FDDAAAG 205
QY 61 VFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALR 120
Db 206 VFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALR 265
QY 121 RGDVWVLLSHDHDGYGAYSNNH 141
Db 266 RGDVWVLLSHDHDGYGAYSNNH 286

RESULT 6

AAB61423
ID AAB61423 standard; protein; 348 AA.
XX
AC AAB61423;
XX
DT 04-APR-2001 (first entry)
XX
DE Human MANGO 245 protein.
XX
KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Homo sapiens.
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -

XX
PS Claim 1; Fig 23; 262pp; English.
XX

CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 348 AA;

Query Match 83.1%; Score 743; DB 22; Length 348;
Best Local Similarity 100.0%; Pred. No. 1.3e-72;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADADAPARGPPAPPEPRSAFSAARTSLVGS DAGPGRHQPLAFDTEFVNIGD FDDAAAG 60
Db 160 ADADAPARGPPAPPEPRSAFSAARTSLVGS DAGPGRHQPLAFDTEFVNIGD FDDAAAG 219

QY 61 VFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALR 120
Db 220 VFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALR 279

QY 121 RGDVWVLLSHDHDGYGAYSNNH 141
Db 280 RGDVWVLLSHDHDGYGAYSNNH 300

RESULT 7

ABG70385
ID ABG70385 standard; Protein; 299 AA.

XX
AC ABG70385;

XX
DT 05-NOV-2002 (first entry)

XX
DE Adipocyte complement-related C1q Tumour Necrosis Factor-like protein #3.

XX
KW Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis;
KW cell signal processing; metabolic pathway modulation; metabolic disorder;
KW obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
KW Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
KW hematopoietic disorder; cirrhosis; pancreatitis; learning defect;
KW memory defect; infertility; congenital heart defect; hair growth;
KW pigmentation disorder; endocrine disorder; respiratory disease; health;
KW gastro-intestinal disease; reproductive; neurological disease;
KW bone marrow transplantation; endocrine disease; allergy; inflammation;
KW nephrological disorder; urinary system disorder; age-related disorder;
KW neuropsychiatric disorder; EGF-related protein; SCUB1; TEN-M4;
KW adipocyte complement-related C1q tumour necrosis factor; out at first;
KW beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;
KW type Ia membrane sushi-containing domain; butyrophillin;
KW type Ia membrane-sushi domain containing.

XX
OS Homo sapiens.

XX
PN WO200257453-A2.

XX
PD 25-JUL-2002.

XX
PF 19-DEC-2001; 2001WO-US50331.

XX
PR 19-DEC-2000; 2000US-265704P.

XX
PR 20-DEC-2000; 2000US-257314P.

XX
PR 02-MAY-2001; 2001US-288153P.

XX
PR 29-MAY-2001; 2001US-294075P.

XX
PR 24-JUL-2001; 2001US-307506P.

XX
PR 10-AUG-2001; 2001US-311590P.

XX
PR 10-AUG-2001; 2001US-311613P.

XX
PR 29-AUG-2001; 2001US-315617P.

PR 14-SEP-2001; 2001US-322358P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
PI Stone DJ, Anderson D, Shimkets RA, Burgess CE, Zerhusen BD, Liu X;
PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
DR WPI; 2002-590744/63.
DR N-PSDB; ABS52097.
XX
XX
PT Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious
PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
PT cancer
XX
XX
PS Claim 1; Page 31; 318bp; English.
XX
XX The present invention relates to new NOVX polypeptides. The invention is
CC useful for treating or preventing a NOVX-associated disorder such as
CC cardiomyopathy or atherosclerosis, where the disorder is related to cell
CC signal processing and metabolic pathway modulation in a subject,
CC preferably human. The invention is also useful for treating metabolic
CC disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
CC disorders, haematopoietic disorders and various cancers. The molecules of
CC the invention are also useful for treating or preventing cirrhosis,
CC pancreatitis, learning and memory defects, infertility, congenital heart
CC defects, acne, hair growth, pigmentation disorders, endocrine disorders,
CC respiratory disease, gastro-intestinal diseases, reproductive, health,
CC neurological diseases, bone marrow transplantation, endocrine diseases,
CC allergy and inflammation, nephrological disorders, urinary system
CC disorders, neuropsychiatric disorders and age-related disorders.
CC The present amino acid sequence represents a NOVX protein of the
CC invention.
XX
XX
SQ Sequence 299 AA;
XX
XX
Query Match 81.7%; Score 730; DB 23; Length 299;
Best Local Similarity 98.6%; Pred. No. 2.9e-71;
Matches 138; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
QY 31 SDAGPGRHQPLAFDTEFVNIGDFFDAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMK 90
Db :|||||
160 ADAGPGRHQPLAFDTEFVNIGDFFDAADVFRCLPGAYFFSFTLGKLPKRTLSVKLMK 219
QY 91 NRDEVQAMITYDDGASRRRQSQSVMLALRRGDVWLISHDHGYGAYSNHGXITFSGF 150
Db :|||||
220 NRDEVQAMITYDDGASRRRQSQSVMLALRRGDVWLISHDHGYGAYSNHGXITFSGF 279
QY 151 LVYPDLPAPAPGGLGASELL 170
Db :|||||
280 LVYPDLPAPAPGGLGASELL 299
XX
XX
RESULT 8
ABG70383
ID ABG70383 standard; Protein; 221 AA.
XX
XX
AC ABG70383;
XX
XX 05-NOV-2002 (first entry)
DT
XX
XX Adipocyte complement-related C1q Tumour Necrosis Factor-like protein #1.
DE
XX
XX Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis;
KW cell signal processing; metabolic pathway modulation; metabolic disorder;
KW obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
KW Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
KW haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
KW memory defect; infertility; congenital heart defect; hair growth;
KW pigmentation disorder; endocrine disorder; respiratory disease; health;
KW gastro-intestinal disease; reproductive; neurological disease;

KW bone marrow transplantation; endocrine disease; allergy; inflammation;
KW nephrological disorder; urinary system disorder; age-related disorder;
KW neuropsychiatric disorder; EGF-related protein; SCUB1; TEN-M4;
KW adipocyte complement-related C1q tumour necrosis factor; out at first;
KW beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;
KW type Ia membrane sushi-containing domain; butyrophillin;
KW
XX
XX Homo sapiens.
XX
XX WO200257453-A2.
XX
XX
PD 25-JUL-2002.
XX
XX
PF 19-DEC-2001; 2001WO-US50331.
XX
XX
PR 19-DEC-2000; 2000US-265704P.
PR 20-DEC-2000; 2000US-257314P.
PR 02-MAY-2001; 2001US-288153P.
PR 29-MAY-2001; 2001US-294075P.
PR 24-JUL-2001; 2001US-307506P.
PR 10-AUG-2001; 2001US-311590P.
PR 10-AUG-2001; 2001US-311613P.
PR 29-AUG-2001; 2001US-315617P.
PR 14-SEP-2001; 2001US-322358P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
PI Stone DJ, Anderson D, Shimkets RA, Burgess CE, Zerhusen BD, Liu X;
PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
XX
DR WPI; 2002-590744/63.
DR N-PSDB; ABS52095.
XX
XX
PT Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious
PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
PT cancer
XX
XX
PS Claim 1; Page 28; 318bp; English.
XX
XX
XX The present invention relates to new NOVX polypeptides. The invention is
CC useful for treating or preventing a NOVX-associated disorder such as
CC cardiomyopathy or atherosclerosis, where the disorder is related to cell
CC signal processing and metabolic pathway modulation in a subject,
CC preferably human. The invention is also useful for treating metabolic
CC disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
CC disorders, haematopoietic disorders and various cancers. The molecules of
CC the invention are also useful for treating or preventing cirrhosis,
CC pancreatitis, learning and memory defects, infertility, congenital heart
CC defects, acne, hair growth, pigmentation disorders, endocrine disorders,
CC respiratory disease, gastro-intestinal diseases, reproductive, health,
CC neurological diseases, bone marrow transplantation, endocrine diseases,
CC allergy and inflammation, nephrological disorders, urinary system
CC disorders, neuropsychiatric disorders and age-related disorders.
CC The present amino acid sequence represents a NOVX protein of the
CC invention.
XX
XX
SQ Sequence 221 AA;
XX
XX
Query Match 77.9%; Score 696; DB 23; Length 221;
Best Local Similarity 76.1%; Pred. No. 1e-67;
Matches 140; Conservative 3; Mismatches 13; Indels 28; Gaps 2;
XX
QY 9 GPPAPP-----EPRSAFSAARTSLVSGDAGPGRHQPLAFDT 46
Db :|||||
44 GPPTPPGLPGRGRAGLSGKNGFPDGSAMRSAFSAARTTPLEGTS-----EMAVTFDK 97
QY 47 EFNIGGDFPDAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNDEVQAMITYDDGASR 106
Db :|||||
98 VYVNIGGDFPDAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNDEVQAMITYDDGASR 157

QY 107 RREMOSQSVMLALRGRDAVWLLSHDHDGYGAYSNHGKYITFSGFLVYPDLAPAPPGIGA 166
Db 158 RREMOSQSVMLALRGRDAVWLLSHDHDGYGAYSNHGKYITFSGFLVYPDLAPAPPGIGA 217
QY 167 SELL 170
Db 218 SELL 221

RESULT 9

AAB61473
ID AAB61473 standard; protein; 134 AA.
AC AAB61473;
XX
DT 04-APR-2001 (first entry)
XX
DE Monkey MANGO 245 C1q domain #2.
XX
XX TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KW pancreatic; skeletal; muscle.
XX
OS Catarhini sp..
XX WO200100672-A1.
PN
XX 04-JAN-2001.
PD
XX 29-JUN-2000; 2000WO-US18184.
PF
XX 29-JUN-1999; 99US-0342687.
PR
XX (MILL-) MILLENNIUM PHARM INC.
PA
XX Holtzman DA, Barnes TM, Fraser CC, Sharp JD;
PI
XX WPI; 2001-050127/06.
DR
XX Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Disclosure; Fig 28; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 134 AA;

Query Match 76.4%; Score 683; DB 22; Length 134;
Best Local Similarity 97.8%; Pred. No. 1.4e-66;
Matches 131; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 19 AFSAAARTSLVGSDDAGPGRHQPLAFDTEFVNIGGDFDAAGVRCRLPGAYFFSFTLGK 78
Db 1 AFSAAARTSLVGSDDAGSGPRHRLAFDTLVLNIGGDFDAAGVRCRLPGAYFFSFTLGK 60
QY 79 LPRKTLVXLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRGRDAVWLLSHDHDGYGAY 138
Db 61 LPRKTLVXLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRGRDAVWLLSHDHDGYGAY 120
QY 139 SNHGKYITFSGFLV 152
Db 121 SNHGKYITFSGFLV 134

RESULT 10
ABG70384
ID ABG70384 standard; Protein; 284 AA.

XX
AC ABG70384;
XX
DT 05-NOV-2002 (first entry)
XX

Adipocyte complement-related C1q Tumour Necrosis Factor-like protein #2.

Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis;
cell signal processing; metabolic pathway modulation; metabolic disorder;
obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
memory defect; infertility; congenital heart defect; hair growth;
pigmentation disorder; endocrine disorder; respiratory disease; health;
gastro-intestinal disease; reproductive; neurological disease;
bone marrow transplantation; endocrine disease; allergy; inflammation;
nephrological disorder; urinary system disorder; age-related disorder;
neuropsychiatric disorder; EGF-related protein; SCUBE1; TEN-M4;
adipocyte complement-related C1q tumour necrosis factor; out at first;
beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;
type 1a membrane sushi-containing domain; butyrophillin;
type 1a membrane-sushi domain containing.

Homo sapiens.

WO200257453-A2.

25-JUL-2002.

19-DEC-2001; 2001WO-US50331.

19-DEC-2000; 2000US-265704P.
20-DEC-2000; 2000US-257314P.
02-MAY-2001; 2001US-288153P.
29-MAY-2001; 2001US-294075P.
24-JUL-2001; 2001US-307506P.
10-AUG-2001; 2001US-311590P.
10-AUG-2001; 2001US-311613P.
29-AUG-2001; 2001US-315617P.
14-SEP-2001; 2001US-322358P.

(CURA-) CURAGEN CORP.

Gangoli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
Stone DJ, Anderson D, Shimkets RA, Burgess CE, Zerhusen BD, Liu X;
Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;

WPI; 2002-590744/63.
N-PSDB; ABSS52096.

Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
atherosclerosis, metabolic disorders, diabetes, obesity, infectious
disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
cancer -

Claim 1; Page 30; 318pp; English.

The present invention relates to new NOVX polypeptides. The invention is
useful for treating or preventing a NOVX-associated disorder such as
cardiomyopathy or atherosclerosis, where the disorder is related to cell
signal processing and metabolic pathway modulation in a subject,
preferably human. The invention is also useful for treating metabolic
disorders (e.g. obesity, diabetes, infectious disease, neurodegenerative
disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
disorders, haematopoietic disorders and various cancers. The molecules of
the invention are also useful for treating or preventing cirrhosis,
pancreatitis, learning and memory defects, infertility, congenital heart
defects, acne, hair growth, pigmentation disorders, endocrine disorders,

CC respiratory disease, gastro-intestinal diseases, reproductive, health,
CC neurological diseases, bone marrow transplantation, endocrine diseases,
CC allergy and inflammation, nephrological disorders, urinary system
CC disorders, neuropsychiatric disorders and age-related disorders.
CC The present amino acid sequence represents a NOVX protein of the
CC invention.

XX
SQ Sequence 284 AA;

Query Match 74.6%; Score 666.5; DB 23; Length 284;
Best Local Similarity 55.4%; Pred. No. 2.3e-64;
Matches 149; Conservative 1; Mismatches 14; Indels 105; Gaps 5;

QY 7 ARGP---PAPPEPSAFSAARTSL-----VGSD----- 32
Db 16 ALGPTPGSGSELRSFAFSAARTPLEGTSEMAVTFDKVYVNIIGDDFVATGQFRCPVGA 75
QY 33 -----AGPGRH-----QPLAF----- 44
Db 76 YFFSFTAGKAPHKSLSVMLVNRDEVQALAFDEQRRPGARRAASQSAMLQLDYGDTWLR 135
QY 45 -----DTEFVNIGGDFDAAGVFCRLPGAYFFSFTLGKLP 81
Db 136 LHGAPQYALGAPGATFSGYLVYADAEFVNIGGDFDAAGVFCRLPGAYFFSFTLGKLP 195
QY 82 KTLVSVKLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRRGDVWLLSHDHGYGAYSNNH 141
Db 196 KTLVSVKLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRRGDVWLLSHDHGYGAYSNNH 255
QY 142 GKTYTFSGFLVYPDLAPAPPGLGASELL 170
Db 256 GKTYTFSGFLVYPDLAPAPPGLGASELL 284

RESULT 11
ABBI1578
ID ABBI1578 standard; peptide; 127 AA.
XX
AC ABBI1578;
XX
DT 11-JAN-2002 (first entry)
DE Human secreted protein homologue, SEQ ID NO:1948.
XX
KM Human; cytokine; cell proliferation; cell differentiation; growth factor;
KM haematopoiesis regulation; tissue growth; immunomodulator; actlyvin;
KM inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KM proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KM myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KM chronic inflammatory condition; proliferative retinopathy;
KM atherosclerosis; coronary heart disease; arterial ischaemia;
KM bone disorder; osteoporosis; vascular growth disorder;
KM tissue regeneration; wound healing; infection; immune disorder;
KM cell culture; drug screening; gene therapy; antiinflammatory;
KM antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KM cyostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KM antifungal; vulneryary; antilulcer.

XX Homo sapiens.
OS
XX
PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03800.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
PI Tang YT, Liu C, Dirmanac RT;
XX

DR WPI; 2001-457740/49.
DR N-PSDB; ABA08822.

PT Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -

XX
PS Claim 20; Page 214-215; 1963pp; English.

CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth factor activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.

XX
SQ Sequence 127 AA;

Query Match 74.3%; Score 664; DB 22; Length 127;
Best Local Similarity 99.2%; Pred. No. 1.6e-64;
Matches 126; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 44 FDTFVNIGGDDPDAAAGVFCRLPGAYFFSFTLGKLPKTLVSVKLMKNRDEVQAMTYDDG 103
Db 1 FDTFVNIGGDDPDAAAGVFCRLPGAYFFSFTLGKLPKTLVSVKLMKNRDEVQAMTYDDG 60
QY 104 ASRRREMOSQSVMLALRRGDVWLLSHDHGYGAYSNNHGKTYTFSGFLVYPDLAPAPPG 163
Db 61 SSRREMOSQSVMLALRRGDVWLLSHDHGYGAYSNNHGKTYTFSGFLVYPDLAPAPPG 120
QY 164 LGASELL 170
Db 121 LGASELL 127

RESULT 12
AAB61470
ID AAB61470 standard; protein; 117 AA.
XX
AC AAB61470;
XX
DT 04-APR-2001 (first entry)

| | | |
|----------|--|---|
| XX | | Human MANGO 245 C1q domain #2. |
| DE | | TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung; |
| XX | | autoimmune; allergy; cardiovascular; brain; degenerative; placental; |
| KW | | pancreatic; skeletal; muscle. |
| KM | | |
| XX | | Homo sapiens. |
| OS | | |
| XX | | WO200100672-A1. |
| PN | | |
| PD | | 04-JAN-2001. |
| XX | | |
| PB | | 29-JUN-2000; 2000WO-US18184. |
| PF | | |
| XX | | 29-JUN-1999; 99US-0342687. |
| PR | | |
| XX | | (MILL-) MILLENNIUM PHARM INC. |
| PA | | |
| XX | | Holtzman DA, Barnes TM, Fraser CC, Sharp JD. |
| PI | | |
| DR | | WPI; 2001-050127/06. |
| XX | | |
| PT | | Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and |
| PT | | MANGO 245 proteins, useful in the treatment of inflammatory diseases |
| PT | | (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver |
| PT | | disorders (e.g. jaundice) - |
| XX | | |
| PS | | Disclosure; Fig 26; 262pp; English. |
| XX | | |
| CC | | The present invention relates to cDNAs encoding TANGO 244, |
| CC | | TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins. |
| CC | | The nucleic acids, proteins and protein modulators are useful for |
| CC | | treating colonic disorders, inflammatory diseases, tumors, |
| CC | | renal disorders, liver disorders, lung disorders, autoimmune diseases, |
| CC | | allergic diseases, cardiovascular diseases, brain disorders, |
| CC | | degenerative diseases placental, pancreatic, skeletal and muscle |
| CC | | disorders. |
| XX | | |
| SQ | Sequence 117 AA; | |
| | Query Match | 68.1%; Score 609; DB 22; Length 117; |
| | Best Local Similarity | 100.0%; Pred. No. 1.5e-58; |
| | Matches 117; Conservative | 0; Mismatches 0; Indels 0; Gaps 0; |
| QY | 19 AFSARTRSLVSGDAGPGPRHQLAFDTFVNIGGDFDAAGVFRCLPGAYFFSFTLGK 78 | |
| DB | 1 AFSARTRSLVSGDAGPGPRHQLAFDTFVNIGGDFDAAGVFRCLPGAYFFSFTLGK 60 | |
| QY | 79 LPRKTLVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLAIRGDVAVWLSDHDGY 135 | |
| DB | 61 LPRKTLVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLAIRGDVAVWLSDHDGY 117 | |
| | RESULT 13 | |
| AAB61468 | AAB61468 standard; protein; 130 AA. | |
| ID | AAB61468 | |
| AC | AAB61468; | |
| XX | | |
| DT | 04-APR-2001 (first entry) | |
| XX | | |
| DE | Mature monkey MANGO 245 protein. | |
| XX | | |
| KW | TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung; | |
| KM | autoimmune; allergy; cardiovascular; brain; degenerative; placental; | |
| KW | pancreatic; skeletal; muscle. | |
| XX | | |
| OS | Catarrhini sp.. | |
| XX | | |
| PN | WO200100672-A1. | |
| XX | | |
| PD | 04-JAN-2001 | |

| | | |
|---------------------------|---|--|
| XX | Pf | -29-JUN-2000; 2000WO-US18184. |
| XX | Pr | -29-JUN-1999; 99US-0342687. |
| XX | Pa | (MILL-) MILLENNIUM PHARM INC. |
| XX | PI | Holtzman DA, Barnes TM, Fraser CC, Sharp JD; |
| XX | DR | WPI; 2001-050127/06. |
| PT | MANGO | Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and |
| PT | TANGO | 245 proteins, useful in the treatment of inflammatory diseases |
| PT | (e.g.) | idiopathic ulcerative colitis), tumors, renal disorders and liver |
| PT | disorders | (e.g. jaundice) - |
| PS | Disclosure: | Page 243; 262pp; English. |
| CC | The present invention | relates to cDNAs encoding TANGO 244, |
| CC | TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins. | |
| CC | The nucleic acids, proteins and protein modulators are useful for | |
| CC | treating colonic disorders, inflammatory diseases, tumors, | |
| CC | renal disorders, liver disorders, lung disorders, autoimmune diseases, | |
| CC | allergic diseases, cardiovascular diseases, brain disorders, | |
| CC | degenerative diseases placental, pancreatic, skeletal and muscle | |
| CC | disorders. | |
| SQ | Sequence | 130 AA; |
| Query Match | 53.9%; Score 482; DB 22; Length 130; | |
| Best Local Similarity | 61.2%; Pred. No. 1.3e-44; | |
| Matches 101; Conservative | 6; Mismatches 22; Indels 36; Gaps 3; | |
| Dc | 6 PARGPAPPEPRSAFSAARTRSLVGSADGPGRRHQPLAFDTFEFNIGDFDAAAGVFCRCR 65 : : : 2 PAPG-PGSSELSAFAARITPLEGAS-----EMAVTFDKYVNIGDFDAATGQFCRCR 54 | |
| Dc | 66 LPGAFFFSFTLGKLPRLTSVKLMKNRDEVQAMIVDDGASRRREMOSGVMLALRRGDVA 125 : : : 55 VPGAFFFSFTVGKAPHKSL-----SVMLALRRGDVA 85 | |
| Dc | 126 WLLSHDHGYGAYSNHGKYITFSGLVYPDLAPAAPGLGASELL 170 86 WLLSHDHGYGAYSNHGKYITFSGLVYPDLAGAAPPGLGAPELL 130 | |
| ID | ABG70386 standard; Protein; 133 AA. | |
| AC | ABG70386; | |
| DT | 05-NOV-2002 (first entry) | |
| DE | Adipocyte complement-related C1q Tumour Necrosis Factor-like protein #4. | |
| Kw | Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis; cell signal processing; metabolic pathway modulation; metabolic disorder; obesity; diabetes; infectious disease; neurodegenerative disorder; acne; Alzheimer's disease; Parkinson's disease; immune disorder; cancer; haematopoietic disorder; cirrhosis; pancreatitis; learning defect; memory defect; infertility; congenital heart defect; hair growth; pigmentation disorder; endocrine disorder; respiratory disease; health; gastro-intestinal disease; reproductive; neurological disease; bone marrow transplantation; endocrine disease; allergy; inflammation; nephrological disorder; urinary system disorder; age-related disorder; neuropsychiatric disorder; EGF-related protein; SCUBEL; TEN-M4; adipocyte complement-related C1q tumour necrosis factor; out at first; beta adrenergic receptor kinase; EphA6/enk-2; glucose transporter; type Ia membrane sushi-containing domain; butyrophilin; type Ia membrane-sushi domain containing. | |
| Os | Homo sapiens. | |

XX
PN WO200257453-A2.
XX
PD 25-JUL-2002.
XX
PF 19-DEC-2001; 2001WO-US50331.
XX
PR 19-DEC-2000; 2000US-265704P.
PR 20-DEC-2000; 2000US-257314P.
PR 02-MAY-2001; 2001US-288153P.
PR 29-MAY-2001; 2001US-294075P.
PR 24-JUL-2001; 2001US-307506P.
PR 10-AUG-2001; 2001US-311590P.
PR 10-AUG-2001; 2001US-311613P.
PR 29-AUG-2001; 2001US-315617P.
PR 14-SEP-2001; 2001US-322358P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Gangoli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R,
PI Stone DJ, Anderson D, Shimkets RA, Burgess CE, Zerhusen BD, Liu X;
PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
DR WPI; 2002-590744/63.
DR N-PSDB; ABS52098.
XX
PT Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious
PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
PT cancer -
XX
PS Claim 1; Page 32; 318pp; English.
XX
CC The present invention relates to new NOVX polypeptides. The invention is
CC useful for treating or preventing a NOVX-associated disorder such as
CC cardiomyopathy or atherosclerosis, where the disorder is related to cell
CC signal processing and metabolic pathway modulation in a subject,
CC preferably human. The invention is also useful for treating metabolic
CC disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
CC disorders, hematopoietic disorders and various cancers. The molecules of
CC the invention are also useful for treating or preventing cirrhosis,
CC pancreatitis, learning and memory defects, infertility, congenital heart
CC defects, acne, hair growth, pigmentation disorders, reproductive health,
CC respiratory disease, gastro-intestinal diseases, endocrine disorders,
CC neurological diseases, bone marrow transplantation, endocrine diseases,
CC allergy and inflammation, nephrological disorders, urinary system
CC disorders, neuropsychiatric disorders and age-related disorders.
CC The present amino acid sequence represents a NOVX protein of the
CC invention.
XX
SQ Sequence 133 AA;

Query Match 48.3%; Score 432; DB 23; Length 133;
Best Local Similarity 56.3%; Pred. No. 3.8e-39;
Matches 94; Conservative 3; Mismatches 18; Indels 52; Gaps 3;

QY 7 ARGP---PAPPEPRSAFSAARTSLVSGDAGPGRHQPLAFDTEFVNIGDFDAAGVFR 63
Db 16 ALGPTPGSGSSELRSAFSAARTTPLEGTS-----EMAVTFDKVYVNIIGDFDAATGRFRC 69

QY 64 CRLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALRRGD 123
Db 70 C-----REMQSQSVMLALRRGD 86

QY 124 AVWLISHDHDGYGAYSNHGKITYTFSGFLVYPDLAPAPPGLGASELL 170
Db 87 AVWLISHDHDGYGAYSNHGKITYTFSGFLVYPDLAPAPPGLGASELL 133

RESULT 15
AAB61488
ID AAB61488 standard; protein; 199 AA.

XX
AC AAB61488;
XX
DT 04-APR-2001 (first entry)
XX
DE Murine MANGO 245 protein.
XX
KM TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
KM autoimmune; allergy; cardiovascular; brain; degenerative; placental;
KM pancreatic; skeletal; muscle.
XX
OS Mus musculus.
XX
PN WO200100672-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US18184.
XX
PR 29-JUN-1999; 99US-0342687.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holzman DA, Barnes TM, Fraser CC, Sharp JD;
XX
DR WPI; 2001-050127/06.
XX
PT Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
PT MANGO 245 proteins, useful in the treatment of inflammatory diseases
PT (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
PT disorders (e.g. jaundice) -
XX
PS Claim 1; Fig 29; 262pp; English.
XX
CC The present invention relates to cDNAs encoding TANGO 244,
CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.
CC The nucleic acids, proteins and protein modulators are useful for
CC treating colonic disorders, inflammatory diseases, tumors,
CC renal disorders, liver disorders, lung disorders, autoimmune diseases,
CC allergic diseases, cardiovascular diseases, brain disorders,
CC degenerative diseases placental, pancreatic, skeletal and muscle
CC disorders.
XX
SQ Sequence 199 AA;

Query Match 40.0%; Score 357.5; DB 22; Length 199;
Best Local Similarity 52.1%; Pred. No. 8.7e-31;
Matches 87; Conservative 10; Mismatches 47; Indels 23; Gaps 6;

QY 7 ARGP---PAPPEPRSAFSAARTSLVSGDAGPGRHQPLAFDTEFVNIGDFDAAGVFR 64
Db 16 ALGPAFGSGSSELRSAFSAARTTPLEGTS-----EMAVTFDKVYVNIIGDFDAATGRFRC 69

QY 65 RLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMITYDDGASRR---REMOSQSVMLALRR 121
Db 70 RVPGAYFFSFTAGKAPHKNLVSVMLVRNDEVQALAFD--KQRRPGARRAASQSAMLQLDY 127

QY 122 GDAVWLISHDHDGYGAYSNHGKITYTFSGFLVYPDL-----APAAP 161
Db 128 GDTVWLRLH---GAPHYALGAPGATFSGYLIVYADADADAPARGPAP 171

Search completed: January 12, 2004, 08:15:36
Job time : 26.6188 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:15:44 ; Search time 20.6986 Seconds
(without alignments)
1655.023 Million cell updates/sec

Title: US-10-085-167-2_COPY_160_358
Perfect score: 894
Sequence: 1 ADADAPARGPPAPPEPRSAF.....LVYPDLAPAPGLGASELL 170

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747907 seqs, 201509753 residues

Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*

10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*

13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*

14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*

15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*

17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------------------|-------------------|
| 1 | 894 | 100.0 | 329 | 12 US-10-236-055A-14 | Sequence 14, Appl |
| 2 | 894 | 100.0 | 329 | 12 US-10-085-167-2 | Sequence 2, Appl |
| 3 | 807.5 | 90.3 | 326 | 12 US-10-236-055A-16 | Sequence 16, Appl |
| 4 | 237.5 | 26.6 | 252 | 10 US-09-893-737-52 | Sequence 52, Appl |
| 5 | 230 | 25.7 | 202 | 15 US-10-091-458-43 | Sequence 43, Appl |
| 6 | 230 | 25.7 | 420 | 12 US-10-236-055A-2 | Sequence 2, Appl |
| 7 | 230 | 25.7 | 746 | 10 US-09-738-973-185 | Sequence 185, App |
| 8 | 230 | 25.7 | 746 | 10 US-09-854-133-185 | Sequence 185, App |
| 9 | 230 | 25.7 | 746 | 15 US-10-144-649A-185 | Sequence 185, App |
| 10 | 230 | 25.7 | 908 | 12 US-10-309-422-10 | Sequence 10, Appl |
| 11 | 230 | 25.7 | 909 | 12 US-10-309-422-22 | Sequence 22, Appl |
| 12 | 230 | 25.7 | 957 | 12 US-10-309-422-14 | Sequence 14, Appl |
| 13 | 230 | 25.7 | 958 | 12 US-10-309-422-26 | Sequence 26, Appl |
| 14 | 230 | 25.7 | 992 | 12 US-10-309-422-18 | Sequence 18, Appl |
| 15 | 230 | 25.7 | 993 | 12 US-10-309-422-30 | Sequence 30, Appl |

| | | | | | |
|----|-------|------|------|----------------------|-------------------|
| 16 | 230 | 25.7 | 1042 | 12 US-10-309-422-8 | Sequence 8, Appl |
| 17 | 230 | 25.7 | 1043 | 12 US-10-309-422-20 | Sequence 20, Appl |
| 18 | 230 | 25.7 | 1091 | 12 US-10-309-422-12 | Sequence 12, Appl |
| 19 | 230 | 25.7 | 1092 | 12 US-10-309-422-24 | Sequence 24, Appl |
| 20 | 230 | 25.7 | 1126 | 12 US-10-309-422-16 | Sequence 16, Appl |
| 21 | 230 | 25.7 | 1127 | 12 US-10-309-422-28 | Sequence 28, Appl |
| 22 | 226 | 25.3 | 409 | 12 US-10-236-055A-4 | Sequence 4, Appl |
| 23 | 220.5 | 24.7 | 281 | 10 US-09-911-176B-44 | Sequence 44, Appl |
| 24 | 220.5 | 24.7 | 281 | 12 US-10-360-186-44 | Sequence 44, Appl |
| 25 | 220.5 | 24.7 | 281 | 12 US-10-392-531-44 | Sequence 44, Appl |
| 26 | 220.5 | 24.7 | 281 | 12 US-10-392-706-44 | Sequence 44, Appl |
| 27 | 220.5 | 24.7 | 281 | 12 US-10-385-015-5 | Sequence 5, Appl |
| 28 | 220.5 | 24.7 | 281 | 15 US-10-180-762-44 | Sequence 44, Appl |
| 29 | 220.5 | 24.7 | 281 | 15 US-10-241-258-44 | Sequence 44, Appl |
| 30 | 208.5 | 23.3 | 198 | 12 US-10-211-689-16 | Sequence 16, Appl |
| 31 | 208.5 | 23.3 | 207 | 12 US-10-211-689-10 | Sequence 10, Appl |
| 32 | 208.5 | 23.3 | 207 | 12 US-10-211-689-12 | Sequence 12, Appl |
| 33 | 208.5 | 23.3 | 214 | 12 US-10-211-689-14 | Sequence 14, Appl |
| 34 | 208.5 | 23.3 | 243 | 9 US-09-866-028-42 | Sequence 42, Appl |
| 35 | 208.5 | 23.3 | 243 | 10 US-09-944-449-42 | Sequence 42, Appl |
| 36 | 208.5 | 23.3 | 243 | 10 US-09-944-457-42 | Sequence 42, Appl |
| 37 | 208.5 | 23.3 | 243 | 10 US-09-944-862-42 | Sequence 42, Appl |
| 38 | 208.5 | 23.3 | 243 | 10 US-09-945-587-42 | Sequence 42, Appl |
| 39 | 208.5 | 23.3 | 243 | 10 US-09-945-015-42 | Sequence 42, Appl |
| 40 | 208.5 | 23.3 | 243 | 10 US-09-944-396-42 | Sequence 42, Appl |
| 41 | 208.5 | 23.3 | 243 | 10 US-09-944-097-42 | Sequence 42, Appl |
| 42 | 208.5 | 23.3 | 243 | 10 US-09-944-432-42 | Sequence 42, Appl |
| 43 | 208.5 | 23.3 | 243 | 10 US-09-943-762-42 | Sequence 42, Appl |
| 44 | 208.5 | 23.3 | 243 | 10 US-09-944-654-42 | Sequence 42, Appl |
| 45 | 208.5 | 23.3 | 243 | 10 US-09-943-851A-42 | Sequence 42, Appl |

ALIGNMENTS

RESULT 1

US-10-236-055A-14

Sequence 14, Application US/10236055A

Publication No. US20030134328A1

GENERAL INFORMATION:

APPLICANT: Basham, Beth E.

APPLICANT: Forsythe, Ian

APPLICANT: Gorman, Daniel M.

APPLICANT: Matsson, Jeanine

APPLICANT: Moshrefi, Mehrdad

APPLICANT: Parham, Christi

TITLE OF INVENTION: MAMMALIAN GENES; RELATED REAGENTS

FILE REFERENCE: DX01343K

CURRENT APPLICATION NUMBER: US/10/236, 055A

CURRENT FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/317, 988

PRIOR FILING DATE: 2001-09-06

NUMBER OF SEQ ID NOS: 36

SOFTWARE: PatentIn version 3.1

SEQ ID NO 14

LENGTH: 329

TYPE: PRT

ORGANISM: Homo sapiens

US-10-236-055A-14

| | | | | |
|-----------------------|-----------------|--------------------------------|------------------------------|-------------|
| Query Match | 100.0% | Score 894; | DB 12; | Length 329; |
| Best Local Similarity | 100.0% | Pred. No. 7.1e-84; | | |
| Matches 170; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| QY | 1 | ADADAPARGPPAPPEPRSAFSAARTSLVGS | DAGGPRHQPLAFDTEFVNIGDFDAAG | 60 |
| DB | 160 | ADADAPARGPPAPPEPRSAFSAARTSLVGS | DAGGPRHQPLAFDTEFVNIGDFDAAG | 219 |
| QY | 61 | VFRCLRGAYFFSFTLGKLPKRLSVKLMKNR | DEVQAMTYDDGASRRREMOSQSVMLALR | 120 |
| DB | 220 | VFRCLRGAYFFSFTLGKLPKRLSVKLMKNR | DEVQAMTYDDGASRRREMOSQSVMLALR | 279 |
| QY | 121 | RGDAVWLISHDHGCGAYSNHGKTYTFSSG | FLVYPDLAPAAPGLGASELL | 170 |

Db
280 RGDVAVLISHDHDGYGAYSNHKKYITFSGFLVYPDLAPAAPGLGASELL 329

RESULT 2

```

US-10-085-167-2
/ Sequence 2, Application US/10085167
/ Publication No. US20030170781A1
/ GENERAL INFORMATION:
/ APPLICANT: Holloway, James L.
/ APPLICANT: Lok, Si
/ TITLE OF INVENTION: SECRETED PROTEIN ZACRPA4
/ FILE REFERENCE: 99-29
/ CURRENT APPLICATION NUMBER: US/10/085,167
/ CURRENT FILING DATE: 2002-02-27
/ PRIOR APPLICATION NUMBER: 60/141,928
/ PRIOR FILING DATE: 1999-07-01
/ NUMBER OF SEQ ID NOS: 9
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 2
/ LENGTH: 329
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-085-167-2

```

| | | | | |
|---------------------------|---------|--------------------|--------|-------------------|
| Query Match | 100.0%; | Score 894; | DB 12; | length 329; |
| Best Local Similarity | 100.0%; | Pred. No. 7.1e-84; | | |
| Matches 170; Conservative | 0; | Mismatches | 0; | Indels 0; Gaps 0; |

QY 1 ADADAPARGPAPPEPRSAFSAARTRSLVGSAGPGGRHOPLAFDTEFFVNIGDFDAAG 60
|||
|||
|||
|||
|||
|||
|||
|||
|||
|||
|||
|||
|||
|||
|||
|||
|||
Dd 160 ADADAPARGPPAPEPRSAFAAARTRSLVGSDAGGPRHQPLAFDTEFFVNIGDFDAAG 219

| | | | | | | | | | |
|----|-----|----------------|------------|----------|--------|----------|----------|-------|-----|
| QY | 61 | VFRCRLPGAYFFSF | FTLGLPKRTL | SVKLMKND | EVQAMI | YDDGASRR | REMQSQSV | MLALR | 120 |
| | | | | | | | | | |
| Db | 220 | VFCRLPGAYFFSF | FTLGLPKRTL | SVKLMKND | EVQAMI | YDDGASRR | REMQSQSV | MLALR | 279 |

| QY | 121 | 280 | 3299 |
|----|--|--|------|
| DB | RGDAVWLISHDHGCGAYSNHGKYYITPSCGLVYPDLAPAAPPGLGASELL | RGDAVWLISHDHGCGAYSNHGKYYITPSCGLVYPDLAPAAPPGLGASELL | |

RESULT 3

```

; Sequence 16, Application US/10236055A
; Publication No. US20030134328A1
;
; GENERAL INFORMATION:
; APPLICANT: Basham, Beth E.
; APPLICANT: Forsythe, Ian
; APPLICANT: Gorman, Daniel M.
; APPLICANT: Mattson, Jeanine
; APPLICANT: Moshrefi, Mehrdad
; APPLICANT: Parham, Christi
; TITLE OF INVENTION: MAMMALIAN GENES; RELATED REAGENTS
; FILE REFERENCE: DX01343K
;
; CURRENT APPLICATION NUMBER: US/10/236,055A
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/317,988
; PRIOR FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Mus musculus
;
US-10-236-055A-16

```

| | | | | |
|-----------------------|-----------------|--------------------|-----------|-------------|
| Query Match | 90.3%; | Score 807.5; | DB 12; | Length 326; |
| Best Local Similarity | 92.9%; | Pred. No. 5.8e-75; | | |
| Matches 157; | Conservative 1; | Mismatches 10; | Indels 1; | Gaps 1; |

QY 1 ADADAPARGPAPPEPRSAFSAARTSLVGSDAGPGPHQPLAFDTEFVNIGDEDDAAG 60

Db 159 ADADAPARG-PAAPEPRSAFSAARTSRSLVGSDAAPGPRHRPLAFDTFLVNIIGDFDAAG 217

Qy 61 VFRCLPGAYFFSFTLGLPKRTISVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLALR 120

Db 218 VFRCLPGAYFFSFTLGLPKRTISVKLMKNRDEVQAMITYDDGASRRREMOSQSVMLPLR 277

RESULT 4

```

US-09-893-737-52
: Sequence 52, Application US/09893737
: Patent No. US20020110855A1
: GENERAL INFORMATION:
:   APPLICANT: Sheppard, Paul O.
:   APPLICANT: Presnell, Scott R.
:   TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
:   FILE REFERENCE: 00-41
:   CURRENT APPLICATION NUMBER: US/09/893, 737
:   CURRENT FILING DATE: 2001-06-28
:   PRIOR APPLICATION NUMBER: US 60/215,446
:   PRIOR FILING DATE: 2000-06-30
:   NUMBER OF SEQ ID NOS: 329
:   SOFTWARE: FastSeq for windows Version 3.0
:   SEQ ID NO 52
:   LENGTH: 252
:   TYPE: PRT
:   ORGANISM: Homo sapiens
US-09-893-737-52

```

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 26.6%; | Score 237.5; | DB 10; | Length 252; |
| Best Local Similarity | 43.0%; | Pred. No. 2.5e-16; | | |
| Matches 65; | Conservative 12; | Mismatches 61; | Indels 13; | Gaps 7; |

QY 9 GPPAPPEPR--SAFSARTRSLVGSADGPGPHQPLAFDTEFVNIGGDFDAAAGVFRCRL 66
 ||| | | | | | |
 : ||| | | | | | :
Db 106 GPPGAACRRAYAFSVGRREGHSSD-----HQAVPFDETELVLNDGAFDLAAGRFLCTV 160

QY 67 PGAFPSFTLGKLPKRTLSVKLMKNDEVOAMIYDDGASRRREMOSQSVMILNRGDVAW 126
 ||| | : ||| : : | ||| : |||
Db 161 PGVFFSLNVHTWNKYETYLHIMLR-RPAVLIAQ-PSEERSVMAQSMLLLAAGDAW 218

QY 127 --LLSHDHDCGYGAYSNHGK-YITFSGFLVYP 154
:
Db 219 VRMFQRDRDN-AIYGEGDL YITFSGHLVYP 248

RESULT 5

US-10-091-458-43
; Sequence 43, Application US/10091458
; Publication No. US20030068627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ06C1
; CURRENT APPLICATION NUMBER: US/10/091,458
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 09/764,900
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26

```

; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
```

```
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08
```

Query Match 25.7%; Score 230; DB 15; Length 202;
Best Local Similarity 38.4%; Pred. No. 1.1e-15;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

```
QY 4 DAPARGP-----PAPPEPSAFSAARTSLVGSADGPGPRHQPLAFDTEFNIGG 53
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 49 DVEVTNPAATILPVHVYPLPQOMRVAFSAART-----SNLAPGTLDPQIVFDLLNNLGE 103
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 54 DFDAAAGVFRCLPGAYFFSFTLGKLP-RKTL SVKLMKNRDEVQAMITYDDGASRRREMOS 112
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 104 TFDLQGRFNCVPVNGTYVFIFHMLKLA VNPPLYVNLKNEEVLVSAYANDGAP-DHETAS 162
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 113 QSVMLALRRGDAVWLLSHDHGYGA-YSNHGKYITFSGFLVYPD 155
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 163 NHAILQLFGDQIWLRLH---RGAIYGSSWKYSTFSGYLLYQD 202
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 6
US-10-236-055A-2

```
; Sequence 2, Application US/10236055A
; Publication No. US20030134328A1
; GENERAL INFORMATION:
; APPLICANT: Basham, Beth E.
; APPLICANT: Forsythe, Ian
; APPLICANT: Gorman, Daniel M.
; APPLICANT: Mattson, Jeanine
; APPLICANT: Moshrefi, Mehrdad
; APPLICANT: Parham, Christi
; TITLE OF INVENTION: MAMMALIAN GENES;-RELATED REAGENTS
; FILE REFERENCE: DX01343K
; CURRENT APPLICATION NUMBER: US/10/236,055A
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/317,988
; PRIOR FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Homo sapiens
```

US-10-236-055A-2

Query Match 25.7%; Score 230; DB 12; Length 420;
Best Local Similarity 38.4%; Pred. No. 2.8e-15;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

```
QY 4 DAPARGP-----PAPPEPSAFSAARTSLVGSADGPGPRHQPLAFDTEFNIGG 53
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 267 DVEVTNPAATILPVHVYPLPQOMRVAFSAART-----SNLAPGTLDPQIVFDLLNNLGE 321
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 54 DFDAAAGVFRCLPGAYFFSFTLGKLP-RKTL SVKLMKNRDEVQAMITYDDGASRRREMOS 112
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 322 TFDLQGRFNCVPVNGTYVFIFHMLKLA VNPPLYVNLKNEEVLVSAYANDGAP-DHETAS 380
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 113 QSVMLALRRGDAVWLLSHDHGYGA-YSNHGKYITFSGFLVYPD 155
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 381 NHAILQLFGDQIWLRLH---RGAIYGSSWKYSTFSGYLLYQD 420
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 7
US-09-738-973-185

```
; Sequence 185, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 185
; LENGTH: 746
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-738-973-185
```

Query Match 25.7%; Score 230; DB 10; Length 746;
Best Local Similarity 38.4%; Pred. No. 5.8e-15;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

```
QY 4 DAPARGP-----PAPPEPSAFSAARTSLVGSADGPGPRHQPLAFDTEFNIGG 53
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 593 DVEVTNPAATILPVHVYPLPQOMRVAFSAART-----SNLAPGTLDPQIVFDLLNNLGE 647
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 54 DFDAAAGVFRCLPGAYFFSFTLGKLP-RKTL SVKLMKNRDEVQAMITYDDGASRRREMOS 112
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 648 TFDLQGRFNCVPVNGTYVFIFHMLKLA VNPPLYVNLKNEEVLVSAYANDGAP-DHETAS 706
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 113 QSVMLALRRGDAVWLLSHDHGYGA-YSNHGKYITFSGFLVYPD 155
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 707 NHAILQLFGDQIWLRLH---RGAIYGSSWKYSTFSGYLLYQD 746
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 8
US-09-854-133-185

```
; Sequence 185, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
```

```
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: Benson, Darin R.
/ APPLICANT: Secrist, Heather
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
/ TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.475C10
/ CURRENT APPLICATION NUMBER: US/09/854,133
/ CURRENT FILING DATE: 2001-05-11
/ NUMBER OF SEQ ID NOS: 735
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 185
/ LENGTH: 746
/ TYPE: PRT
/ ORGANISM: Homo sapien
US-09-854-133-185
```

```
Query Match      25.7%; Score 230; DB 10; Length 746;
Best Local Similarity 38.4%; Pred. No. 5.8e-15;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;
```

```
QY 4 DAPARGP-----PAPPEPSAFSAARTSLVGS DAPGPRHQPLAFDTEFVNIGG 53
Db 593 DVPVTNPAATILPVHVYPLPQOMRVA FSAART-----SNLAPGTLDPQPIVFDLLNNLGE 647

QY 54 DFDAAAGVFCRLPGAYFFSFTLGKLP-RKTL SVKLMKNRDEVQAMITYDDGASRRREMOS 112
Db 648 TFDLQLGRFNCVPVNGTYVFIFHMLKLA VNVPLVYVNLKNEEVLVSAYANDGAP-DHETAS 706

QY 113 QSVMLALRRGDAVWLLSHDHDGYGA-YSNHGKYITFSGFLVYPD 155
Db 707 NHAIIQLFGQDIWLRH---RGAIVGSSWKYSTFGYLLYQD 746
```

```
RESULT 9
US-10-144-649A-185
/ Sequence 185, Application US/10144649A
/ Publication No. US20030118599A1
/ GENERAL INFORMATION:
/ APPLICANT: Lodes, Michael J.
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Fan, Liqun
/ APPLICANT: Algate, Paul A.
/ APPLICANT: McNeill, Patricia D.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
/ TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.475C11
/ CURRENT APPLICATION NUMBER: US/10/144,649A
/ CURRENT FILING DATE: 2002-08-21
/ NUMBER OF SEQ ID NOS: 749
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 185
/ LENGTH: 746
/ TYPE: PRT
/ ORGANISM: Homo sapien
US-10-144-649A-185
```

```
Query Match      25.7%; Score 230; DB 15; Length 746;
Best Local Similarity 38.4%; Pred. No. 5.8e-15;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;
```

```
QY 4 DAPARGP-----PAPPEPSAFSAARTSLVGS DAPGPRHQPLAFDTEFVNIGG 53
Db 593 DVPVTNPAATILPVHVYPLPQOMRVA FSAART-----SNLAPGTLDPQPIVFDLLNNLGE 647

QY 54 DFDAAAGVFCRLPGAYFFSFTLGKLP-RKTL SVKLMKNRDEVQAMITYDDGASRRREMOS 112
Db 648 TFDLQLGRFNCVPVNGTYVFIFHMLKLA VNVPLVYVNLKNEEVLVSAYANDGAP-DHETAS 706

QY 113 QSVMLALRRGDAVWLLSHDHDGYGA-YSNHGKYITFSGFLVYPD 155
Db 707 NHAIIQLFGQDIWLRH---RGAIVGSSWKYSTFGYLLYQD 746
```

```
RESULT 10
US-10-309-422-10
/ Sequence 10, Application US/10309422
/ Publication No. US20030139587A1
/ GENERAL INFORMATION:
/ APPLICANT: Walke, D. Wade
/ APPLICANT: Wilganowski, Nathaniel L.
/ APPLICANT: Turner, C. Alexander Jr.
/ TITLE OF INVENTION: NO. US20030139587A1e1 Human Proteins and Polynucleotides Encodin
/ FILE REFERENCE: LEX-0142-USA
/ CURRENT APPLICATION NUMBER: US/10/309,422
/ CURRENT FILING DATE: 2002-12-03
/ PRIOR APPLICATION NUMBER: US/09/798,771
/ PRIOR FILING DATE: 2001-03-02
/ PRIOR APPLICATION NUMBER: US 60/186,557
/ PRIOR FILING DATE: 2000-03-02
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 10
/ LENGTH: 908
/ TYPE: PRT
/ ORGANISM: homo sapiens
US-10-309-422-10
```

```
Query Match      25.7%; Score 230; DB 12; Length 908;
Best Local Similarity 38.4%; Pred. No. 7.4e-15;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;
```

```
QY 4 DAPARGP-----PAPPEPSAFSAARTSLVGS DAPGPRHQPLAFDTEFVNIGG 53
Db 755 DVPVTNPAATILPVHVYPLPQOMRVA FSAART-----SNLAPGTLDPQPIVFDLLNNLGE 809

QY 54 DFDAAAGVFCRLPGAYFFSFTLGKLP-RKTL SVKLMKNRDEVQAMITYDDGASRRREMOS 112
Db 810 TFDLQLGRFNCVPVNGTYVFIFHMLKLA VNVPLVYVNLKNEEVLVSAYANDGAP-DHETAS 868

QY 113 QSVMLALRRGDAVWLLSHDHDGYGA-YSNHGKYITFSGFLVYPD 155
Db 869 NHAIIQLFGQDIWLRH---RGAIVGSSWKYSTFGYLLYQD 908
```

```
RESULT 11
US-10-309-422-22
/ Sequence 22, Application US/10309422
/ Publication No. US20030139587A1
/ GENERAL INFORMATION:
/ APPLICANT: Walke, D. Wade
/ APPLICANT: Wilganowski, Nathaniel L.
/ APPLICANT: Turner, C. Alexander Jr.
/ TITLE OF INVENTION: NO. US20030139587A1e1 Human Proteins and Polynucleotides Encodin
/ FILE REFERENCE: LEX-0142-USA
/ CURRENT APPLICATION NUMBER: US/10/309,422
/ CURRENT FILING DATE: 2002-12-03
/ PRIOR APPLICATION NUMBER: US/09/798,771
/ PRIOR FILING DATE: 2001-03-02
/ PRIOR APPLICATION NUMBER: US 60/186,557
/ PRIOR FILING DATE: 2000-03-02
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 22
/ LENGTH: 909
/ TYPE: PRT
/ ORGANISM: homo sapiens
US-10-309-422-22
```

```
Query Match      25.7%; Score 230; DB 12; Length 909;
Best Local Similarity 38.4%; Pred. No. 7.5e-15;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;
```

```
QY 4 DAPARGP-----PAPPEPSAFSAARTSLVGS DAPGPRHQPLAFDTEFVNIGG 53
Db 756 DVPVTNPAATILPVHVYPLPQOMRVA FSAART-----SNLAPGTLDPQPIVFDLLNNLGE 810
```



```

QY      54 DFDAAAGVFCRLPGAYFFSFYTLGKL-P-RTLSVKLKNRDEVQAMITYDDGASRRREMOS 112
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      811 TFDLQLGRFNCVPNGTYVFI FHMLKLVNVPLVYVNLMKNEEVLVSAYANDGAP-DHETAS 869
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      113 QSVMLALRGGDAVWLLSHDHGYG-YSNHGKYITFSGFLVYPD 155
      ::|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      870 NHAIIQLFGDDQIWLRLH---RGATYGSSWKYSTFSGYLLYQD 909
      ::|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

RESULT 12
US-10-309-422-14
; Sequence 14, Application US/10309422
; Publication No. US20030139587A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20030139587A1 Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: LEX-0142-USA
; CURRENT APPLICATION NUMBER: US/10/309,422
; PRIOR FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: US/09/798,771
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,557
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 957
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-309-422-14

```

| Query Match | 25.7%; | Score 230; | DB 12; | Length 957; |
|-----------------------|--------|--|----------------|--------------------|
| Best Local Similarity | 38.4%; | Pred. No. 8e-15; | | |
| Matches | 63; | Conservative 16; | Mismatches 63; | Indels 22; Gaps 6; |
| QY | 4 | DAPARGP-----PAPPEPRSAFSAARTRSLVSGDAGGPRHQPPLAFDTEFVNIGG | 53 | |
| DB | 804 | DVPVTNPAAITILPVHVHYPLPQQMRVAFSAART-----SNLAPGTLTDQPIVFEDLLNLTIGE | 858 | |
| QY | 54 | DFDAAGVFRCLPFGAYFFSFTLGKLP-RKTL SVKLMKNRDEVQAMITYDDGASRRREMOS | 112 | |
| DB | 859 | TFDLQLGRFNCPVNGTYVIFHMLKLA VNVPLVYVNLMKNEEVLVSAYANDGAP-DHETAS | 917 | |
| QY | 113 | QSVMLALRRGDAVWMLSHDHDGYGA-YSNHGKYITFSGFIVYPD | 155 | |
| DB | 918 | NHAIIQLFGQDQIWLRLH---RGAIIYSSSWKYSTFSGYLLYOD | 957 | |

```

RESULT 13
US-10-309-422-26
; Sequence 26, Application US/10309422
; Publication No. US20030139587A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20030139587A1el Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: LEX-0142-USA
; CURRENT APPLICATION NUMBER: US/10/309,422
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: US/09/798,771
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,557
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 958
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-309-422-26

```

```

Query Match      25.7%; Score 230; DB 12; Length 958;
Best Local Similarity 38.4%; Pred. No. 8e-15;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

QY      4 DAPARGP-----PAPPEPRSAFSARTRSLVGS DAGPGPRHQPLAFDTEFVNIG 53
      | | | | | | | | | | | | | | | | | | | | | | | |
Db      805 DVPTNPAAATILPVHVYPLPQOMRVAFSAART-----SNLAPGTLDPQIVFDLLNNLGE 859

QY      54 DFDAAGVFRGRLPGAYFFSFTLGKLPRKRTLSVKLMKNRDEVQAMITYDDGASRRREMOS 112
      | | | | | | | | | | | | | | | | | | | | | | | |
Db      860 TFDLQLGRFNCVNGTYVFI FHLMLKLA VNVPLVYVNLMKNEEVLVSAYANDGAP-DHETAS 918

QY      113 QSVMLALRRGDVAVWLLSHDHDGYGA-YSNHGKYITFSGFLVYPD 155
      : | | | | | | | | | | | | | | | | | | | | | | | |
Db      919 NHAITQLFGQDQIWLRLH---RGAITYGSSWKYSTFSGYLLLYOD 958

```

```

RESULT 14
US-10-309-422-18
; Sequence 18, Application US/10309422
; Publication No. US20030139587A1
; GENERAL INFORMATION:
; APPLICANT: walke, D. Wade
; APPLICANT: wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20030139587A1e1 Human Proteins and Polynucleotides Encodin
; FILE REFERENCE: LEX-0142-USA
; CURRENT APPLICATION NUMBER: US/10/309,422
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: US/09/798,771
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,557
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 18
; LENGTH: 992
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-309-422-18

```

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 25.7%; | Score 230; | DB 12; | Length 992; |
| Best Local Similarity | 38.4%; | Pred. No. 8.3e-15; | | |
| Matches 63; | Conservative 16; | Mismatches 63; | Indels 22; | Gaps 6; |

| | | | | | | |
|----|-----|----------------------|---------------------|---------------------|----------------------|-------------------------|
| QY | 4 | DAPARGP----- | PAPPEPRSAFSARTSLVGS | DAGPGRHOPLAFDTEFVNI | G | 53 |
| | | | : | : | : | |
| Db | 839 | DVPVTNPAAITLIPVHVYPL | PQOMRVAFA | SART----- | SNLAPGTLDPQIVFDLLNNL | GE 893 |
| QY | 54 | DFDAAGVFRCLP | GAYFFSFTL | GKL | P-RKTL | SVKLMKNRDE |
| | | | | | | EQAMITDDGASRR |
| Db | 894 | TFDLQLGRFNC | PVNGTYVFIF | FMMLKLA | VNVPLVYNLMKNE | EVLSAYANDGAP-DHETAS 952 |
| | | | : | : | : | |
| QY | 113 | QSVMLALRRGDA | VWLLSHD | HDGYGA-YSNH | KYITFS | GFLVYPD 155 |
| | | : | | | | |
| Db | 953 | NHAILQLFGDQI | WLRLH--- | RGAITYGSSW | KYSTFS | GYYLLYQD 992 |
| | | : | | | | |

```

RESULT 15
US-10-309-422-30
; Sequence 30, Application US/10309422
; Publication No. US20030139587A1
; GENERAL INFORMATION:
; APPLICANT: walke, D. Wade
; APPLICANT: wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20030139587A1 Human Proteins and Polynucleotides Encodir
; FILE REFERENCE: LEX-0142-USA
; CURRENT APPLICATION NUMBER: US/10/309,422
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: US/09/798,771
; PRIOR FILING DATE: 2001-03-02

```

```

; PRIOR APPLICATION NUMBER: US 60/186,557
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 993
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-309-422-30

```

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 25.7%; | Score 230; | DB 12; | Length 993; |
| Best Local Similarity | 38.4%; | Pred. No. 8.3e-15; | | |
| Matches 63; | Conservative 16; | Mismatches 63; | Indels 22; | Gaps 6; |

```

QY      1 DAPARGP-----PAPPEPRSAFSAARTRSLVSGDAGPGPRHOPLAFDTFEVNIIG 53
Db      840 DVPVTPNPAATILPVHVYPLPQOMRVAFSAART-----SNLAPGTLDPIVFDLLNNLGE 894
QY      54 DFDAAAGVFCRLPGAYFFSFTLGKL-P-RKTLVSKLMKNRDEVQAMIVDDGASRRREMQS.112
Db      895 TFDLQLGRFNCVPVNGTYVFI FHM LKLANVPLVYNLMKNEEVLVSAYANDGAP-DHETAS 953
QY      113 QSVMLALRRGDVAVMLLSHDHDGYGA-YSNHGKYITFSGLVYPPD 155
Db      954 NHAIIQLLFQGDQIWLRLH----RGAIYGSWSWKYSTFSGYLLYOD 993

```

Search completed: January 12, 2004, 08:33:37
Job time : 20.6986 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:14 ; Search time 8.65269 Seconds
(without alignments)
831.284 Million cell updates/sec

Title: US-10-085-167-2_COPY_160_358
Perfect score: 894
Sequence: 1 ADADAPARGPPAPPEPRSAF.....LVYPDLAPAPGIGASELL 170

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------------------|-------------------|
| 1 | 230 | 25.7 | 746 | 4 US-09-370-838-185 | Sequence 185, App |
| 2 | 220.5 | 24.7 | 281 | 3 US-09-118-408-44 | Sequence 44, Appl |
| 3 | 220.5 | 24.7 | 281 | 4 US-09-506-855-44 | Sequence 44, Appl |
| 4 | 220.5 | 24.7 | 281 | 4 US-09-911-176B-44 | Sequence 44, Appl |
| 5 | 220.5 | 24.7 | 281 | 4 US-09-619-740-44 | Sequence 44, Appl |
| 6 | 220.5 | 24.7 | 281 | 4 US-09-506-852-44 | Sequence 44, Appl |
| 7 | 208.5 | 23.3 | 228 | 4 US-09-336-536-4 | Sequence 4, Appli |
| 8 | 208.5 | 23.3 | 243 | 3 US-09-140-804-2 | Sequence 2, Appli |
| 9 | 208.5 | 23.3 | 243 | 4 US-09-336-536-3 | Sequence 3, Appli |
| 10 | 208.5 | 23.3 | 243 | 4 US-09-686-838B-2 | Sequence 2, Appli |
| 11 | 202.5 | 22.7 | 281 | 3 US-09-118-408-2 | Sequence 2, Appli |
| 12 | 202.5 | 22.7 | 281 | 4 US-09-506-855-2 | Sequence 2, Appli |
| 13 | 202.5 | 22.7 | 281 | 4 US-09-911-176B-2 | Sequence 2, Appli |
| 14 | 202.5 | 22.7 | 281 | 4 US-09-619-740-2 | Sequence 2, Appli |
| 15 | 202.5 | 22.7 | 281 | 4 US-09-506-852-2 | Sequence 2, Appli |
| 16 | 199.5 | 22.3 | 243 | 3 US-09-188-930-295 | Sequence 295, App |
| 17 | 199.5 | 22.3 | 243 | 4 US-09-312-283C-295 | Sequence 295, App |
| 18 | 194.5 | 21.8 | 228 | 4 US-09-336-536-11 | Sequence 11, Appl |
| 19 | 194.5 | 21.8 | 243 | 4 US-09-336-536-10 | Sequence 10, Appl |
| 20 | 194 | 21.7 | 259 | 4 US-09-996-243-47 | Sequence 47, Appl |
| 21 | 194 | 21.7 | 260 | 4 US-09-489-847-198 | Sequence 198, App |
| 22 | 194 | 21.7 | 287 | 4 US-09-489-847-349 | Sequence 349, App |
| 23 | 188 | 21.0 | 247 | 2 US-08-463-911-2 | Sequence 2, Appli |
| 24 | 188 | 21.0 | 247 | 4 US-09-776-976-4 | Sequence 4, Appli |
| 25 | 187.5 | 21.0 | 247 | 4 US-09-909-547-4 | Sequence 4, Appli |
| 26 | 187.5 | 21.0 | 294 | 3 US-09-188-930-294 | Sequence 294, App |
| 27 | 187.5 | 21.0 | 294 | 4 US-09-312-283C-294 | Sequence 294, App |

| | | | | | |
|----|-------|------|-----|----------------------|-------------------|
| 28 | 187 | 20.9 | 247 | 3 US-09-140-804-8 | Sequence 8, Appli |
| 29 | 187 | 20.9 | 247 | 3 US-09-118-408-3 | Sequence 3, Appli |
| 30 | 187 | 20.9 | 247 | 4 US-09-506-855-3 | Sequence 3, Appli |
| 31 | 187 | 20.9 | 247 | 4 US-09-686-838B-8 | Sequence 8, Appli |
| 32 | 187 | 20.9 | 247 | 4 US-09-911-176B-3 | Sequence 3, Appli |
| 33 | 187 | 20.9 | 247 | 4 US-09-619-740-3 | Sequence 3, Appli |
| 34 | 187 | 20.9 | 247 | 4 US-09-506-852-3 | Sequence 3, Appli |
| 35 | 184.5 | 20.6 | 285 | 4 US-09-312-283C-382 | Sequence 382, App |
| 36 | 182 | 20.4 | 247 | 4 US-09-776-976-2 | Sequence 2, Appli |
| 37 | 182 | 20.4 | 247 | 4 US-09-909-547-2 | Sequence 2, Appli |
| 38 | 177.5 | 19.9 | 215 | 3 US-09-140-804-5 | Sequence 5, Appli |
| 39 | 177.5 | 19.9 | 215 | 4 US-09-686-838B-5 | Sequence 5, Appli |
| 40 | 177.5 | 19.9 | 215 | 4 US-09-911-176B-50 | Sequence 50, Appl |
| 41 | 177.5 | 19.9 | 215 | 4 US-09-619-740-53 | Sequence 53, Appl |
| 42 | 175 | 19.6 | 231 | 4 US-09-530-423-2 | Sequence 2, Appli |
| 43 | 175 | 19.6 | 244 | 2 US-08-463-911-7 | Sequence 7, Appli |
| 44 | 175 | 19.6 | 244 | 3 US-09-140-804-3 | Sequence 3, Appli |
| 45 | 175 | 19.6 | 244 | 4 US-09-336-536-20 | Sequence 20, Appl |

ALIGNMENTS

RESULT 1
US-09-370-838-185
; Sequence 185, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370, 838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 185
; LENGTH: 746
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-370-838-185

Query Match 25.7%; Score 230; DB 4; Length 746;
Best local Similarity 38.4%; Pred. No. 2.6e-17;
Matches 63; Conservative 16; Mismatches 63; Indels 22; Gaps 6;

QY 4 DAPARGP-----PAPPEPRSAFSAARTSLVSGDAGPGRHOPLAFDTEFVNIGG 53
DB 593 DVPVTNPATILPVHVYPLPQGMRYAFSAART-----SNLAPGTLDQPIVFDDLNLNIGE 647

QY 54 DFDAAAGVRCRLPGAYFFSFTLGKLP-RKTLVSKLMKNRDEVQAMITYDGASRRREMOS 112
DB 648 TFDLQGRFNCVPNGTYVFIFHMLKLVNVPVLYNLMKNEEVLVSAYANDGAP-DHETAS 706

QY 113 QSVMLALRRGDAVWLSHDHGYGCA-YSNHGKYITFSGFVLYPD 155
DB 707 NHAIIQLFGQDIWLRH----RGAIYGSSWKYSTFSGYLYQD 746

RESULT 2
US-09-118-408-44
; Sequence 44, Application US/09118408A
; Patent No. 6265544
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
; FILE REFERENCE: 97-30
; CURRENT APPLICATION NUMBER: US/09/118,408A


```
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/053,154
; EARLIER FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 44
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-118-408-44
```

```
Query Match      24.7%; Score 220.5; DB 3; Length 281;
Best Local Similarity 39.2%; Pred. No. 8.1e-17;
Matches 56; Conservative 20; Mismatches 58; Indels 9; Gaps 4;
```

```
QY 18 SAFSAARTSLVGS D A G P G R H O P L A F D T E F V N I G D F D A A G V F R C R L P G A Y F F S F T L G 77
Db 146 AAFSVGRKKALHSND-----YFQPVVFDTEFVNLYKHFMFTGKFYCYVPGIYFFSLNVH 200
QY 78 KLPRKTL SVKLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRRGDVW--LLSHDHGY 135
Db 201 TWNQKETYLHIMKNEEV-VILYAQ-VSDRSIMOSQSLMELREDEDEVWVRLFKGERENA 258
QY 136 GAYSNHGKIYTFSGFLVY PDLAP 158
Db 259 IFSDEFDTYITFSGYLVK PASEP 281
```

RESULT 3

```
US-09-506-855-44
; Sequence 44, Application US/09506855
; Patent No. 6448221
```

```
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Lasser, Gerald W.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: INHIBITORS FOR USE IN HEMOSTASIS AND
; TITLE OF INVENTION: IMMUNE FUNCTION
; FILE REFERENCE: 99-12
; CURRENT APPLICATION NUMBER: US/09/506,855
; CURRENT FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 44
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-506-855-44
```

```
Query Match      24.7%; Score 220.5; DB 4; Length 281;
Best Local Similarity 39.2%; Pred. No. 8.1e-17;
Matches 56; Conservative 20; Mismatches 58; Indels 9; Gaps 4;
```

```
QY 18 SAFSAARTSLVGS D A G P G R H O P L A F D T E F V N I G D F D A A G V F R C R L P G A Y F F S F T L G 77
Db 146 AAFSVGRKKALHSND-----YFQPVVFDTEFVNLYKHFMFTGKFYCYVPGIYFFSLNVH 200
QY 78 KLPRKTL SVKLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRRGDVW--LLSHDHGY 135
Db 201 TWNQKETYLHIMKNEEV-VILYAQ-VSDRSIMOSQSLMELREDEDEVWVRLFKGERENA 258
QY 136 GAYSNHGKIYTFSGFLVY PDLAP 158
Db 259 IFSDEFDTYITFSGYLVK PASEP 281
```

RESULT 4

```
US-09-911-176B-44
; Sequence 44, Application US/0991176B
; Patent No. 6518403
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: ANTIBODIES THAT BIND AN
```

```
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOG
; FILE REFERENCE: 97-30D1
; CURRENT APPLICATION NUMBER: US/09/911,176B
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: 09/118,408
; PRIOR FILING DATE: 1998-07-17
; PRIOR APPLICATION NUMBER: 60/053,154
; PRIOR FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 44
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-911-176B-44
```

```
Query Match      24.7%; Score 220.5; DB 4; Length 281;
Best Local Similarity 39.2%; Pred. No. 8.1e-17;
Matches 56; Conservative 20; Mismatches 58; Indels 9; Gaps 4;
```

```
QY 18 SAFSAARTSLVGS D A G P G R H O P L A F D T E F V N I G D F D A A G V F R C R L P G A Y F F S F T L G 77
Db 146 AAFSVGRKKALHSND-----YFQPVVFDTEFVNLYKHFMFTGKFYCYVPGIYFFSLNVH 200
QY 78 KLPRKTL SVKLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRRGDVW--LLSHDHGY 135
Db 201 TWNQKETYLHIMKNEEV-VILYAQ-VSDRSIMOSQSLMELREDEDEVWVRLFKGERENA 258
QY 136 GAYSNHGKIYTFSGFLVY PDLAP 158
Db 259 IFSDEFDTYITFSGYLVK PASEP 281
```

RESULT 5

```
US-09-619-740-44
; Sequence 44, Application US/09619740
; Patent No. 6544946
```

```
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Lasser, Gerald W.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: INHIBITORS FOR USE IN HEMOSTASIS AND IMMUNE FUNCTION
; FILE REFERENCE: 99-12C3
; CURRENT APPLICATION NUMBER: US/09/619,740
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/253,604
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/444,794
; PRIOR FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: 09/506,855
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 44
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-619-740-44
```

```
Query Match      24.7%; Score 220.5; DB 4; Length 281;
Best Local Similarity 39.2%; Pred. No. 8.1e-17;
Matches 56; Conservative 20; Mismatches 58; Indels 9; Gaps 4;
```

```
QY 18 SAFSAARTSLVGS D A G P G R H O P L A F D T E F V N I G D F D A A G V F R C R L P G A Y F F S F T L G 77
Db 146 AAFSVGRKKALHSND-----YFQPVVFDTEFVNLYKHFMFTGKFYCYVPGIYFFSLNVH 200
QY 78 KLPRKTL SVKLMKNRDEVQAMTYDDGASRRREMOSQSVMLALRRGDVW--LLSHDHGY 135
Db 201 TWNQKETYLHIMKNEEV-VILYAQ-VSDRSIMOSQSLMELREDEDEVWVRLFKGERENA 258
QY 136 GAYSNHGKIYTFSGFLVY PDLAP 158
```

Db 259 IFSDEFDITYTFSGYLKPASEP 281

RESULT 6

```

US-09-506-852-44
; Sequence 44, Application US/09506852
; Patent No. 6566499
; GENERAL INFORMATION:
; APPLICANT: Shepard, Paul O.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
; FILE REFERENCE: 97-30
; CURRENT APPLICATION NUMBER: US/09/506, 852
; CURRENT FILING DATE: 2000-02-17
; EARLIER APPLICATION NUMBER: 60/053,154
; EARLIER FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 44
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-506-852-44

```

| | | | | |
|-----------------------|--------|--------------------|--------|----------------|
| Query Match | 24.7%; | Score 220.5; | DB 4; | Length 281; |
| Best Local Similarity | 39.2%; | Pred. No. 8.1e-17; | | |
| Matches | 56; | Conservative. | 20; | Mismatches 58; |
| | | | Indels | 9; |
| | | | Gaps | 4; |

18 SAFSARTSLVGS D A G P G P R H Q P L A F D T E F V N I G G D F D A A G V F R C R L P G A Y F F S F T L G 77

Db 146 AAFSVGRKKALHSND-----YEQPVVFDTFVNLYKHENMTGKFCYVPGIYFFSLNVH 200

QY 78 KLPRKTL SVKLMKNRDEVOAMITYDDGASRRREMOSQSVMLALRRGDVW--LLSHDPDGY 135

Db 201 TWNQKETYLHIMNEEV-VILYAQ-VSDRSIMQSQSLMELREDEDEVTWVRLFKGERENA 258

QY 136 GAYSNHGKYITFSGFLVYPDLAP 158

Db 259 IFSDEFDITYTFSGYLKPPASEP 281

RESULT 7

```

US-09-336-536-4
; Sequence 4, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-336-536-4

```

| | | | | |
|-----------------------|--------|--------------------|--------|----------------|
| Query Match | 23.3%; | Score 208.5; | DB 4; | Length 228; |
| Best Local Similarity | 35.7%; | Pred. No. 1.4e-15; | | |
| Matches | 56; | Conservative | 18; | Mismatches 72; |
| | | | Indels | 11; |
| | | | Gaps | 3; |

QY 4 DAPARGPAPP-----EPRSAFSARTRSLVGS D A G P G P R H O P L A F D T E F V N I G G D F D A 58

Db 70 EAGPAGPTGPAGECSVPPPSAFAFAKRSSESRV-----PPPSDAPLPFDRVLVNEQGHYDAV 124

QY 59 AGVFRCLPGAYFESFTLGKLPKRTLSVKLMKNRDEVQAMLYDDGASRRREMOSQSVMLA 118

Db 125 T G K F T C Q V P G V Y F A - V H A T V Y R A S L Q F D L V K N G E S I A S F F Q F E G G W P K P A S L S G G A M V R 183

QY 119 LRRGDVWLLSHDHDGYGAYSNHGK YITFSGELVYPD 155

Db 184 LEPEDQVWVGVDYIGIYASIKTDTSTFSGFLVYSD 220

RESULT 8

```

US-09-140-804-2
; Sequence 2, Application US/09140804
; Patent No. 6197930
; GENERAL INFORMATION:
; APPLICANT: Shepard, Paul O.
; APPLICANT: Humes, Jacqueline M.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
; FILE REFERENCE: 97-49
; CURRENT APPLICATION NUMBER: US/09/140,804
; CURRENT FILING DATE: 1998-08-26
; EARLIER APPLICATION NUMBER: 60/056,983
; EARLIER FILING DATE: 1997-08-26
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 2
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-140-804-2

```

| | | | | |
|-----------------------|--------|--------------------|-------|----------------|
| Query Match | 23.3% | Score 208.5; | DB 3; | Length 243; |
| Best Local Similarity | 35.7%; | Pred. No. 1.5e-15; | | |
| Matches | 56; | Conservative | 18; | Mismatches 72; |
| | | | | Indels 11; |
| | | | | Gaps 3; |

QY 4 DAPARGPAPP-----EPRSAFSAARTRSLVCS DAGPGPRHQPLAFDTEFVNIGGDFDAA 58

Db 85 EAGPAGPTGAGECSVPPRSAFSAKRSESR-----PPSDAPLPFDRVLVNEQGHYDAV 139

59 AGVFRCLPGAYFESFTLGKLPKRTLSVKLMKNRDEVQAMIYDDGASRRREMOSQSVMLA 118

Db 140 T G K F T C Q V P G V Y Y F A - V H A T V Y R A S L Q F D L V K N G E S I A S F Q F F G G W P K P A S L S G G A M V R 198

QY 119 LRRGDAVWLSDHDCYGAYSNHGYITFSGFLVYPD 155

Db 199 LEPEDOVWVGVCVDYIGTYASTKTDSTESGETWSPD 235

RESULT 9

```

US-09-336-536-3
; Sequence 3, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Boscone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-336-536-3

```

| | | | | |
|-----------------------|--------|--------------------|-------|----------------|
| Query Match | 23.3%; | Score 208.5; | DB 4; | Length 243; |
| Best Local Similarity | 35.7%; | Pred. No. 1.5e-15; | | |
| Matches | 56; | Conservative | 18; | Mismatches 72; |
| | | | | Indels 11; |
| | | | | Gaps 3; |

4 DAPARGPPAP-----EPRSAFSAARTRSLVGS D A G P G P R H Q P L A F D T E F V N I G G D F D A 58

Db 85 EAGPAGPTGPAGECSVPPRSAFSAKRSESRV-----PPPSDAPLPEDRVLVNEQGHYDAV 139

59 AGVFCRLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMIYDDGASRRREMQSQSVMLA 118 QY

Db 140 TGKFTCGVPGVYFA-VHATVYRASLQFDLVKNGESIASFFQFFGGWPKPASLSGAMVR 198


```
Matches    60;   Conservative    18;   Mismatches    58;   Indels    33;   Gaps    6;
QY          5 APARGPPAP-----PEPR-----SAFSAARTSLVGS D A G P G R H -----QPLAFD 45
           | ||| | | | | | | | | : ||| | | : | |
DB          119 AGARGHGTGPKGQKSGMGA FGERCKSHYAFAFSVGKK-----PMHSNHYYQT V I F D 168
              | : | | | | | | | | : | | | | : | |
QY          46 TEFVNIGGD EDAAGVFRCRLPGA YFFSFTLGKLPRKTLSVKLMKNRDEVQA MIYDDGAS 105
           ||||| : | : | | | | | | : | | | | : | |
DB          169 TEFVNLYDHFMFTGKFYCYPGLYFFSLNVHTWNQKETYLHIMKNEEEVVILFAQVG-- 226
              | : | | | | | | | | : | : | | | | : | |
QY          106 RRREMOSQSVM LALRRGD AVLISHDHDGYGAYS NH--GKYITFGSEFLV 152
           | ||||| : | | | | | | : | : | | | | : | |
DB          227 DRSIMOSQLMELREQDVWVRPLYKGERENAI FSEELDTYTITFGSYLV 275
              | ||||| : | | | | | | : | : | | | | : | |

RESULT 14
US-09-619-740-2
; Sequence 2, Application US/09619740
; Patent No. 6544946
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Lasser, Gerald W.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: INHIBITORS FOR USE IN HEMOSTASIS AND IMMUNE FUNCTION
; FILE REFERENCE: 99-12C3
; CURRENT APPLICATION NUMBER: US/09/619,740
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/253,604
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 09/444,794
; PRIOR FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: 09/506,855
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-619-740-2
```

```

Query Match      22.7%; Score 202.5; DB 4; Length 281;
Best Local Similarity 35.5%; Pred. No. 8.7e-15;
Matches 60; Conservative 18; Mismatches 58; Indels 33; Gaps 6;

QY      5 APARGPPAP-----PEPR-----SAFSAARTSLVGS DAGPGRH-----QPIAFD 45
          ||| | | | | | | | : ||| | | : ||| | | |
DB     119 AGARGHGTGPKGKGSMGAPGERCKSHYAFAFSVGKK-----PMHSNHYQTVAIFD 168
          ||| | | | | | | | : ||| | | : ||| | | |
QY      46 TEFNIGGDFDAAGVFRCLPGAYFFSFTLGKLPKRTLSVKLMKNRDEVQAMTYDDGAS 105
          ||||| : : : ||| | | | | : : : ||| : ||| |
DB     169 TEFVNLYDHFMFTGKFYCYPGLYFFSLNVHTWNOKETYLHIMKNEEAVITLFAQVG-- 226
          ||||| : : : ||| | | | | : : : ||| : ||| |
QY      106 RRREMOSQSVMLALRRGDVWLSSHHDGYGAYSNN--GKYITFSGFLV 152
          + ||||| : ||| | | | | : : : : : ||||| : |||
DB     227 DRSIMOSQLMLETREQDVWRLYKGERENAI FSEELDTYITTFSGYLV 275
          ||||| : ||| | | | | : : : : : ||||| : |||

RESULT 15
US-09-506-852-2
; Sequence 2, Application US/09506852
; Patent No. 6566499
; GENERAL INFORMATION:
; APPLICANT: Shepard, Paul O.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
; FILE REFERENCE: 97-30
; CURRENT APPLICATION NUMBER: US/09/506,852
; CURRENT FILING DATE: 2000-02-17
; EARLIER APPLICATION NUMBER: 60/053,154
; EARLIER FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2

```

```

; LENGTH: 281
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-506-852-2

```

| | | | | |
|-----------------------|--------|--------------------|--------|----------------|
| Query Match | 22.7%; | Score 202.5; | DB 4; | Length 281; |
| Best Local Similarity | 35.5%; | Pred. No. 8.7e-15; | | |
| Matches | 60; | Conservative | 18; | Mismatches 58; |
| | | | Indels | 33; |
| | | | Gaps | 6; |

```

QY      5  APARGPPAP-----PEPR-----SAFSAARTSLVGSADAGCPGRH-----QPLAFD  45
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db     119  AGARGHTGPKGKGSMGAPGERCKSHYAAFSVGRKK-----PMHSNHYYQTIVFD  168
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY      46  TEFVNIGGDFDAAGVFRCLPGAYFFSFTLGLPRKTLVTKLMKNRDEYQAMITYDDGAS  105
      ||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db     169  TEFVNLXDFNMFTGKFYCYVPGLYFFSLNVHTWNQKETYLHIMKNEEEVILLFAQVG--  226
      ||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY      106  RRREMOSQSVMLALRRGDAVWLSHDHGCGAYSNH--GKYITFSGLV  152
      ||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db     227  DRSIMOSQSLMLELRQDQVAVRLYKGERENAI FSEELDTYITFSGLV  275
      ||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

```

Search completed: January 12, 2004, 08:20:05
Job time : 9.65269 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:09 ; Search time 1.79441 Seconds
(without alignments)
1661.397 Million cell updates/sec

Title: US-10-085-167-3
Perfect score: 57
Sequence: 1 FXXXXXXXXXXXXXXXXXXXXXGXVFX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Match | Query Length | DB ID | Description |
|------------|-------|---------|--------------|----------|--------------------|
| 1 | 22 | 38.6 | 250 | 2 E83651 | hypothetical prote |
| 2 | 22 | 38.6 | 302 | 2 D82112 | tyra protein VC214 |
| 3 | 22 | 38.6 | 421 | 2 A99309 | membrane transport |
| 4 | 22 | 38.6 | 423 | 2 H90490 | membrane transport |
| 5 | 22 | 38.6 | 441 | 2 T19306 | hypothetical prote |
| 6 | 22 | 38.6 | 868 | 2 T02635 | D2 protein homolog |
| 7 | 22 | 38.6 | 3705 | 2 AD0123 | probable autotrans |
| 8 | 22 | 38.6 | 4976 | 2 T14165 | peptide synthetase |
| 9 | 21 | 36.8 | 156 | 2 E72052 | rRNA methylase - C |
| 10 | 21 | 36.8 | 156 | 2 A86573 | rRNA methylase - C |
| 11 | 21 | 36.8 | 210 | 2 F64609 | conserved hypochet |
| 12 | 21 | 36.8 | 210 | 2 B71906 | hypothetical prote |
| 13 | 21 | 36.8 | 212 | 2 H81344 | hypothetical prote |
| 14 | 21 | 36.8 | 231 | 2 B83060 | hypothetical prote |
| 15 | 21 | 36.8 | 238 | 2 I64038 | hypothetical prote |
| 16 | 21 | 36.8 | 269 | 2 T49293 | hypothetical prote |
| 17 | 21 | 36.8 | 277 | 2 T06760 | hypothetical prote |
| 18 | 21 | 36.8 | 303 | 2 T35616 | probable membrane |
| 19 | 21 | 36.8 | 303 | 2 A12543 | hypothetical prote |
| 20 | 21 | 36.8 | 307 | 2 A71602 | rifin PFB0955w - m |
| 21 | 21 | 36.8 | 325 | 2 T31989 | hypothetical prote |
| 22 | 21 | 36.8 | 420 | 2 AG3057 | glucose-1-phosphat |
| 23 | 21 | 36.8 | 420 | 2 G98228 | glucose-1-phosphat |
| 24 | 21 | 36.8 | 450 | 2 B96561 | hypothetical prote |
| 25 | 21 | 36.8 | 508 | 2 T04605 | hypothetical prote |
| 26 | 21 | 36.8 | 516 | 2 C82946 | hypothetical prote |
| 27 | 21 | 36.8 | 555 | 2 T01391 | WD-repeat protein |
| 28 | 21 | 36.8 | 653 | 2 T03319 | gene 112 protein - |
| 29 | 21 | 36.8 | 669 | 2 T24571 | hypothetical prote |

| | | | | | |
|----|----|------|------|----------|--------------------|
| 30 | 21 | 36.8 | 704 | 2 T13665 | NADH2 dehydrogenas |
| 31 | 21 | 36.8 | 713 | 2 AE0531 | lysine decarboxyla |
| 32 | 21 | 36.8 | 724 | 2 A42953 | nitrous oxide redu |
| 33 | 21 | 36.8 | 724 | 2 D96949 | secreted protein c |
| 34 | 21 | 36.8 | 779 | 2 T44659 | nitrous oxide redu |
| 35 | 21 | 36.8 | 826 | 2 E81706 | conserved hypochet |
| 36 | 21 | 36.8 | 885 | 2 AC2444 | hypothetical prote |
| 37 | 21 | 36.8 | 976 | 2 G82209 | GDH family prote |
| 38 | 21 | 36.8 | 1138 | 2 G71554 | probable transmemb |
| 39 | 21 | 36.8 | 1316 | 2 B86240 | protein F20B24.12 |
| 40 | 21 | 36.8 | 1417 | 2 H90670 | probable invasins |
| 41 | 21 | 36.8 | 1417 | 2 D85521 | probable adhesin |
| 42 | 20 | 35.1 | 40 | 2 G81731 | hypothetical prote |
| 43 | 20 | 35.1 | 119 | 2 A53257 | H+-transporting AT |
| 44 | 20 | 35.1 | 127 | 2 H64012 | sufi protein homol |
| 45 | 20 | 35.1 | 129 | 2 B90175 | NADH dehydrogenase |

ALIGNMENTS

RESULT 1
E83651
hypothetical protein BH0013 [imported] - Bacillus halodurans (strain C-125)
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C/Accession: E83651
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: E83651
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-250 <STO>
A/Cross-references: GB:AP001507; GB:BA000004; NID:g10172612; PIDN:BAB03732.1; GSPDB:GN001
A/Experimental source: strain C-125
C/Genetics:
A/Gene: BH0013

Query Match 38.6%; Score 22; DB 2; Length 250;
Best Local Similarity 36.4%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 19 FXXXXXXXXGXFX 29
Db 132 FTFSGAGSYDF 142

RESULT 2
D82112
tyra protein VC2145 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C/Species: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C/Accession: D82112
R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A/Reference number: A82035; MUID:20406833; PMID:10952301
A/Accession: D82112
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-302 <HEI>

A/Cross-references: GB:AE004287; GB:AE003852; NID:g9656689; PIDN:AAF95290.1; GSPDB:GN001;
A/Experimental source: serogroup O1; strain N16961; biotype El Tor
C/Genetics:
A/Gene: VC2145
A/Map position: 1
C/Superfamily: Escherichia coli hypothetical protein b2431

Query Match 38.6%; Score 22; DB 2; Length 302;

Best Local Similarity 36.4%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXGXYYF 29

DB 277 FTKAVTGAYFF 287

RESULT 3

A99309

membrane transporter SSO1505 [imported] - Sulfolobus solfataricus

C/Species: Sulfolobus solfataricus

C/Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C/Accession: A99309

R/She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.
arrett, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A/Description: Sulfolobus solfataricus complete genome.

A/Reference number: A99139

A/Accession: A99309

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-421 <KUR>

A/Cross-references: GB:AE006641; NID:gl3814734; PIDN:AAK41728.1; GSPDB:GN00155

C/Genetics:

A/Gene: SSO1505

Query Match

Best Local Similarity 36.4%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXGXYYF 29

DB 338 FAIAAIGYSF 348

RESULT 4

H90490

membrane transporter SSO3079 [imported] - Sulfolobus solfataricus

C/Species: Sulfolobus solfataricus

C/Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C/Accession: H90490

R/She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.
arrett, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A/Description: Sulfolobus solfataricus complete genome.

A/Reference number: A99139

A/Accession: H90490

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-423 <KUR>

A/Cross-references: GB:AE006641; NID:gl3816494; PIDN:AAK43183.1; GSPDB:GN00155

C/Genetics:

A/Gene: SSO3079

Query Match

Best Local Similarity 36.4%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXGXYYF 29

DB 340 FAIAAIGYSF 350

RESULT 5

T19306

hypothetical protein C15C8.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C/Accession: T19306

R/Sims, M.

submitted to the EMBL Data Library, June 1996

A/Reference number: Z19106

A/Accession: T19306

A/Status: preliminary; translated from GB/EMBL/DBDJ

A/Molecule type: DNA

A/Residues: 1-441 <WIL>

A/Cross-references: EMBL:Z75527; PIDN:CAA99774.1; GSPDB:GN00023; CESP:C15C8.1

A/Experimental source: clone C15C8

C/Genetics:

A/Gene: CESP:C15C8.1

A/Map position: 5

A/Introns: 156/2; 205/3; 258/1

Query Match

Best Local Similarity 36.4%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXGXYYF 29

DB 165 FAAILSGSYHF 175

RESULT 6

T02635

D2 protein homolog - slime mold (Dictyostelium discoideum) plasmid Ddps

C/Species: Dictyostelium discoideum

C/Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 24-Nov-1999

C/Accession: T02635

R/Rieben Jr., W.K.; Gonzales, C.M.; Gonzales, S.T.; Pilkington, K.J.; Kiyosawa, H.; Hughe
Genetics 148, 1117-1125, 1998

A/Title: Dictyostelium discoideum nuclear plasmid Ddps is a chimera related to the Ddpl

A/Reference number: Z14684; MUID:98198836; PMID:9539429

A/Accession: T02635

A/Status: preliminary; translated from GB/EMBL/DBDJ

A/Molecule type: DNA

A/Residues: 1-868 <RIE>

A/Cross-references: EMBL:AF000580; NID:g3068582; PIDN:AAK14391.1; PID:g3068584

A/Experimental source: strain WS2162

C/Genetics:

A/Gene: d2

A/Genome: plasmid

A/Mobile element: plasmid Ddps

C/Superfamily: slime mold (Dictyostelium discoideum) plasmid Ddps D2 protein homolog

Query Match

Best Local Similarity 36.4%; Pred. No. 7.2e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXGXYYF 29

DB 345 FTTYXSGFYSP 355

RESULT 7

AD0123

probable autotransporter protein yapH [imported] - Yersinia pestis (strain CO92)

C/Species: Yersinia pestis

C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001

C/Accession: AD0123

R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; F
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, F
Nature 413, 523-527, 2001

A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A/Reference number: AB0001; MUID:21470413; PMID:11586360

A/Accession: AD0123

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-3705 <KUR>

A/Cross-references: GB:AL590842; PIDN:CAC89847.1; PID:g15979073; GSPDB:GN00175

C/Genetics:

A/Gene: yapH

```
Query Match      38.6%; Score 22; DB 2; Length 3705;
Best Local Similarity 36.4%; Pred. No. 3.3e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      19 FXXXXXXGXXXP 29
Db      3553 FTASVEGGYAF 3563

RESULT 8
T14165
peptide synthetase homolog - Mycobacterium smegmatis
C:Species: Mycobacterium smegmatis
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 01-Dec-2000
C:Accession: T14165
R:Yu, S.; Fiss, E.; Jacobs Jr., W.R.
J. Bacteriol. 180, 4676-4685, 1998
A:Title: Analysis of the exochelin locus in mycobacterium smegmatis: biosynthesis genes
A:Reference number: Z17898; MUID:98389687; PMID:9721311
A:Accession: T14165
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-4976 <YUS>
A:Cross-references: EMBL:AF027770; NID:g3560502; PID:g3560507; PIDN:AAC82550.1
C:Genetics:
A:Gene: fxbC
C:Superfamily: acetate-CoA ligase homology; acyl carrier protein homology
C:Keywords: carrier protein; phosphopantetheine; phosphoprotein
F:68-566/Domain: acetate-CoA ligase homology <ACL>
F:582-649/Domain: acyl carrier protein homology <ACPI>
F:1620-2062/Domain: acetate-CoA ligase homology <ACLI>
F:2078-2144/Domain: acyl carrier protein homology <ACP2>
F:3139-3591/Domain: acetate-CoA ligase homology <ACL2>
F:3608-3679/Domain: acyl carrier protein homology <ACP3>
F:4198-4620/Domain: acetate-CoA ligase homology <ACL3>
F:4637-4705/Domain: acyl carrier protein homology <ACP4>
F:614,3643,4669/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match      38.6%; Score 22; DB 2; Length 4976;
Best Local Similarity 36.4%; Pred. No. 4.5e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      19 FXXXXXXGXXXP 29
Db      1738 FAGTSGLYDF 1748

RESULT 9
E72052
rRNA methylase - Chlamydomophila pneumoniae (strain CWL029)
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 03-Nov-2000
C:Accession: E72052
R:Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Clamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: E72052
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-156 <ARN>
A:Cross-references: GB:AE001648; GB:AE001363; NID:g4376946; PIDN:AAD18799.1; PID:g437695
A:Experimental source: strain CWL029
C:Genetics:
A:Gene: spoU_2
C:Superfamily: Chlamydomophila pneumoniae rRNA methylase

Query Match      36.8%; Score 21; DB 2; Length 156;
Best Local Similarity 36.4%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      19 FXXXXXXGXXXP 29
Db      1 FXXXXXXGXXXP 29
```

```
Db      88 FSLPSSGTYVF 98

RESULT 10
A86573
rRNA methylase [imported] - Chlamydomophila pneumoniae (strain J138)
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: A86573
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: A86573
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-156 <STO>
A:Cross-references: GB:BA000008; NID:g8979032; PIDN:BAA98867.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: spoU_2
C:Superfamily: Chlamydomophila pneumoniae rRNA methylase

Query Match      36.8%; Score 21; DB 2; Length 156;
Best Local Similarity 36.4%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      19 FXXXXXXGXXXP 29
Db      88 FSLPSSGTYVF 98

RESULT 11
F64609
conserved hypothetical integral membrane protein HP0718 - Helicobacter pylori (strain 26k
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 17-Mar-2000
C:Accession: F64609
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Gilek, A.; McKenne
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wathey, L.
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.A
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: F64609
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-210 <TOM>
A:Cross-references: GB:AE000585; GB:AE000511; NID:g2313845; PIDN:AAD07771.1; PID:g231384
C:Superfamily: Escherichia coli hypothetical 23k protein (sbm-fba intergenic region)

Query Match      36.8%; Score 21; DB 2; Length 210;
Best Local Similarity 36.4%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      19 FXXXXXXGXXXP 29
Db      75 FGALFTGYAF 85

RESULT 12
B71906
hypothetical protein jhp0656 - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 17-Mar-2000
C:Accession: B71906
R:Alm, R.A.; Ling, L.S.L.; Molr, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682
```


A/Accession: B71906
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-210 <ARN>
A/Cross-references: GB:AE001497; GB:AE001439; NID:g4155199; PIDN:AAD06232.1; PID:g415520
A/Experimental source: strain J99
C/Genetics:
A/Gene: jhp0656
C/Superfamily: Escherichia coli hypothetical 23k protein (sbm-fba intergenic region)

Query Match 36.8%; Score 21; DB 2; Length 210;
Best Local Similarity 36.4%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29
DB 75 FGAVFTGFYAF 85

RESULT 13
H81344
hypothetical protein Cj0733 [imported] - Campylobacter jejuni (strain NCTC 11168)
C/Species: Campylobacter jejuni
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C/Accession: H81344
R/Parkhill, J.; Wren, B.W.; Mungall, K.; Kettle, J.M.; Churcher, C.; Basham, D.; Chillin
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrel
Nature 403, 665-668, 2000
A/Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A/Reference number: A81250; MUID:20150912; PMID:10688204
A/Accession: H81344
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-212 <PAR>
A/Cross-references: GB:AL139076; GB:AL111168; NID:g6968128; PIDN:CAB73007.1; PID:g696819
A/Experimental source: serotype O2, strain NCTC 11168
C/Genetics:
A/Gene: Cj0733

Query Match 36.8%; Score 21; DB 2; Length 212;
Best Local Similarity 36.4%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29
DB 137 FSGKAYGYSF 147

RESULT 14
B83060
hypothetical protein PA4685 [imported] - Pseudomonas aeruginosa (strain PA01)
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C/Accession: B83060
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bz
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
. ; LoY, S.; Olson, M.V.
Nature 406, 959-964, 2000
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A/Reference number: A82950; MUID:20437337; PMID:10984043
A/Accession: B83060
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-231 <STO>
A/Cross-references: GB:AE004882; GB:AE004091; NID:g9950939; PIDN:AAG08072.1; GSPDB:GN001
A/Experimental source: strain PA01
C/Genetics:
A/Gene: PA4685

Query Match 36.8%; Score 21; DB 2; Length 231;
Best Local Similarity 36.4%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29
DB 146 FASEDSGYVRF 156

RESULT 15
I64038
hypothetical protein H1626 - Haemophilus influenzae (strain Rd KW20)
C/Species: Haemophilus influenzae
C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 08-Oct-1999
C/Accession: I64038
R/Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
. D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A/Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, C
A/Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A/Reference number: A64000; MUID:95350630; PMID:7542800
A/Accession: I64038
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-238 <TIGR>
A/Cross-references: GB:U32835; GB:L42023; NID:g1574459; PIDN:AAC23269.1; PID:g1574467; T]

Query Match 36.8%; Score 21; DB 2; Length 238;
Best Local Similarity 36.4%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29
DB 122 FALAGMSYF 132

Search completed: January 12, 2004, 08:16:52
Job time : 4.79441 secs


```
RUN3_MOUSE
ID RUN3_MOUSE STANDARD; PRT; 409 AA.
AC Q64131; Q99P92; Q9R199;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Runt-related transcription factor 3 (Core-binding factor, alpha 3 subunit) (CBF-alpha 3) (Acute myeloid leukemia 2 protein) (Oncogene AML-2) (Polyomavirus enhancer binding protein 2 alpha C subunit) (PEBP2-alpha C) (PEA2-alpha C) (SL3-3 enhancer factor 1 alpha C subunit) (SL3/AKV core-binding factor alpha C subunit).
DE subunit) (SL3/AKV core-binding factor alpha C subunit).
DE RUNX3 OR CBF3 OR AML2 OR PEBP2A3.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Spleen;
RA Levanon D., Negreanu V., Groner Y.;
RT "The mouse AML2/Cbf3 cdna sequence.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 92-167 FROM N.A.
RX MEDLINE=95331802; PubMed=7607690;
RA Wijemanga C., Speck N.A., Dracopoli N.C., Hofker M.H., Liu P., Collins F.S.;
RT "Identification of a new murine runt domain-containing gene, Cbf3, and localization of the human homolog, CBF3, to chromosome 1p35-pter.";
RL Genomics 26:611-614(1995).
RN [3]
RP SEQUENCE OF 1-95 FROM N.A. (ISOFORM 1).
RX MEDLINE=95331801; PubMed=7607689;
RA Calabi F., Rhodes M., Williamson P., Boyd Y.;
RT "Identification and chromosomal mapping of a third mouse runt-like locus.";
RL Genomics 26:607-610(1995).
RN [4]
RP SEQUENCE OF 1-95 FROM N.A. (ISOFORM 1).
RC STRAIN=129/Sv;
RA Negreanu V., Levanon D., Bettoun J.D., Groner Y.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP PARTIAL SEQUENCE FROM N.A. (ISOFORM 2).
RA Bangsow C., Rubins N., Bernstein Y.;
RT "RUNX3 gene structure and function.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: BINDS TO THE CORE SITE OF MURINE LEUKEMIA VIRUS, THE CORE SEQUENCES IN THE ENHANCER OF THE POLYOMAVIRUS, AND ALSO TO THE ENHANCERS OF THE T-CELL RECEPTOR GENES. MAY BE INVOLVED IN THE CONTROL OF CELLULAR PROLIFERATION AND/OR DIFFERENTIATION (BY SIMILARITY).
CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit. The alpha subunit binds DNA as a monomer and through the runt domain. DNA-binding is increased by heterodimerization. Interacts with TLE1 (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q64131-1; Sequence=Displayed;
CC IsoId=Q64131-2; Sequence=VSP_005950;
CC -1- DOMAIN: A proline/serine/threonine rich region at the C-terminus is necessary for transcriptional activation of target genes.
CC -1- SIMILARITY: Contains 1 Runt domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF155880; AAD38985.1; -.
DR EMBL; S78518; AAB34843.1; -.
DR EMBL; AF169246; AAD46381.1; -.
DR EMBL; AF321443; AAK1181.1; -.
DR PIR; A56842; A56842.
DR HSSP; O60472; 1CWO.
DR MGD; MGI:102672; Runt3.
DR GO; GO:0007411; Paxon guidance; IMP.
DR InterPro; IPR000040; AML1_Runt.
DR Pfam; PF00853; Runt; 1.
DR PRINTS; PR00967; ONCOGENEAML1.
KW Transcription regulation; DNA-binding; Nuclear protein; ATP-binding;
KW Alternative splicing.
FT DOMAIN 23 28 POLY-GLY.
FT DOMAIN 55 182 RUNT.
FT NP BIND 143 150 ATP (POTENTIAL).
FT VARSPPLIC 1 5 MRIPV -> MASNSIFDSFPQLYTNLHT (in isoform 2).
SQ SEQUENCE 409 AA; 43628 MW; 83440C67FC591008 CRC64;
/FTId=VSP_005950.
Query Match 38.6%; Score 22; DB 1; Length 409;
Best Local Similarity 36.4%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 19 FXXXXXGXXF 29
Db 331 FYGASSGSYQF 341
-----
RESULT 3
Y995_THETN STANDARD; PRT; 175 AA.
ID Y995_THETN
AC Q6R330;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein TTE0995.
GN TTE0995.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J., Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L., Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
CC -1- SIMILARITY: BELONGS TO THE UPF0189 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE013064; AAM24250.1; -.
DR HAMAP; MF_01205; -; 1.
DR InterPro; IPR002589; Alpp.
DR Pfam; PF01661; Alpp; 1.
DR SMART; SM00506; Alpp; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 175 AA; 18762 MW; C1FB4A03A8A1B5A CRC64;
```

```
Query Match          36.8%; Score 21; DB 1; Length 175;
Best Local Similarity 36.4%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY      19 FXXXXXXGXXXF 29
      118 FPSISTGAYGF 128

Db

RESULT 4
Y177_METMA          STANDARD; PRT; 187 AA.
ID      Y177_METMA
AC      O8Q0F9;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DE      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Hypothetical protein MM0177.
GN      MM0177.
OS      Methanosarcina mazel (Methanosarcina frisia).
OC      Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC      Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX      NCBI_TaxID=2209;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Goel / G01 / ATCC BAA-199 / DSM 3647 / OCM 88;
RX      MEDLINE=22120827; PubMed=12125824;
RA      Deppenmeier U., Johann A., Hartsech T., Merkl R., Schmitz R.A.,
RA      Martinez-Arias R., Henne A., Wierzer A., Baumer S., Jacobi C.,
RA      Brueggemann H., Lienard T., Christmann A., Boemecke M., Steckel S.,
RA      Bhattacharya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA      Fritz H.-J., Gottschalk G.;
RT      "The genome of Methanosarcina mazel: evidence for lateral gene
RT      transfer between Bacteria and Archaea.";
RL      J. Mol. Microbiol. Biotechnol. 4:453-461 (2002).
CC      -!- SIMILARITY: BELONGS TO THE UPF0189 FAMILY.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      -----
DR      EMBL; AE013241; AAM29873.1; -.
DR      HAMAP; MF_01205; -; 1.
DR      InterPro; IPR002589; Alpp.
DR      Pfam; PF01661; Alpp; 1.
KW      Hypothetical protein; Complete proteome.
SQ      SEQUENCE 187 AA; 20361 MW; 8E125281358A1F62 CRC64;

Query Match          36.8%; Score 21; DB 1; Length 187;
Best Local Similarity 36.4%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY      19 FXXXXXXGXXXF 29
      130 FPAISTGAYGF 140

Db

RESULT 5
YG14_METAC          STANDARD; PRT; 195 AA.
ID      YG14_METAC
AC      O8TOD0;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Hypothetical protein MA1614.
GN      MA1614.
OS      Methanosarcina acetivorans.
OC      Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC      Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX      NCBI_TaxID=2214;
CC
```

```
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C2A / ATCC 35395 / DSM 2834;
RX      MEDLINE=21929760; PubMed=11932238;
RA      Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA      FitzHugh W., Calvo S., Engels R., Smirnov S., Alnoor D., Brown A.,
RA      Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
RA      Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA      Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA      Hedderich R., Ingram-Smith C., Kuettnner H.C., Krzycki J.A.,
RA      Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA      Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA      Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA      Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA      Metcalf W.W., Birren B.;
RT      "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT      and physiological diversity.";
RL      Genome Res. 12:532-542 (2002).
CC      -!- SIMILARITY: BELONGS TO THE UPF0189 FAMILY.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      -----
DR      EMBL; AE010833; AAM05027.1; -.
DR      HAMAP; MF_01205; -; 1.
DR      InterPro; IPR002589; Alpp.
DR      Pfam; PF01661; Alpp; 1.
DR      SMART; SMO0506; Alpp; 1.
KW      Hypothetical protein; Complete proteome.
SQ      SEQUENCE 195 AA; 21671 MW; 35DCEA8EDB0C34A0 CRC64;

Query Match          36.8%; Score 21; DB 1; Length 195;
Best Local Similarity 36.4%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY      19 FXXXXXXGXXXF 29
      138 FPTISTGAYGF 148

Db

RESULT 6
P24_CRIGR           STANDARD; PRT; 196 AA.
ID      P24_CRIGR
AC      P49020;
DT      01-FEB-1996 (Rel. 33, Created)
DT      01-FEB-1996 (Rel. 33, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Cop-coated vesicle membrane protein p24 precursor (Fragment).
GN      RNP24.
OS      Cricetulus griseus (Chinese hamster).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC      Cricetulus.
OX      NCBI_TaxID=10029;
RN      [1]
RP      SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX      MEDLINE=95372409; PubMed=7644530;
RA      Stames M.A., Craighead M.W., Hoe M.H., Lampen N., Geromanos S.,
RA      Tempst P., Rothman J.E.;
RT      "An integral membrane component of coatomer-coated transport vesicles
RT      defines a family of proteins involved in budding.";
RL      Proc. Natl. Acad. Sci. U.S.A. 92:8011-8015 (1995).
CC      -!- FUNCTION: Could have a role in the budding of coatomer-coated and
CC      other species of coated vesicles. Could bind cargo molecules to
CC      collect them into budding vesicles.
CC      -!- SUBUNIT: Interacts with ARFGAP1, inhibiting the GTPase-activating
CC      activity of ARFGAP1 (By similarity).
CC      -!- SUBCELLULAR LOCATION: Type I membrane protein. Golgi-derived
```


CC coatomer-coated vesicles.
CC -1- SIMILARITY: BELONGS TO THE EMP24/GP25L FAMILY.
CC -1- SIMILARITY: Contains 1 GOLD domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U26264; AAA82925.1; -.
DR InterPro; IPR000348; Emp24_gp25L_p24.
DR Pfam; PF01105; EMP24_GP25L; 1.
DR PROSITE; PS50866; GOLD; 1.
KM Transport; Protein transport; Transmembrane; Signal; Golgi stack.
FT NON TER 1 1
FT SIGNAL <1 15 POTENTIAL.
FT CHAIN 16 196 COP-COATED VESICLE MEMBRANE PROTEIN P24.
FT DOMAIN 16 163 LUMENAL (POTENTIAL).
FT TRANSMEM 164 184 POTENTIAL.
FT DOMAIN 185 196 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 107 GOLD.
SQ SEQUENCE 196 AA; 22189 MW; 63502CF101B77810 CRC64;

Query Match 36.8%; Score 21; DB 1; Length 196;
Best Local Similarity 36.4%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXXF 29
| | | | |
Db 77 FAAMDGTYKF 87

RESULT 7
YB05_THEAC
ID YB05_THEAC STANDARD; PRT; 196 AA.
AC Q9HJ67;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein Tai105.
GN Tai105.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; Pubmed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum.";
RL Nature 407:508-513(2000).
CC -1- SIMILARITY: BELONGS TO THE UPF0189 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL445066; CAC12232.1; ALT_INIT.
DR HAMAP; MF_01205; -; 1.
DR InterPro; IPR002589; A1pp.
DR Pfam; PF01661; A1pp; 1.
DR SMART; SM00506; A1pp; 1.
KM Hypothetical protein; Complete proteome.

SQ SEQUENCE 196 AA; 21664 MW; 06F0D91D97665BE2 CRC64;

Query Match 36.8%; Score 21; DB 1; Length 196;
Best Local Similarity 36.4%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXXF 29
| | | | |
Db 126 FPALSTGAYGF 136

RESULT 8
P24_HUMAN
ID P24_HUMAN STANDARD; PRT; 201 AA.
AC Q15363;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cop-coated vesicle membrane protein p24 precursor (p24A).
GN RNP24.
OS Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96291865; Pubmed=8663407;
RA Blum R., Feick P., Puype M., Vandekerckhove J., Klengel R.,
RA Nastainczyk W., Schulz I.;
RT "Timp21 and p24A, two type I proteins enriched in pancreatic
RT microsomal membranes, are members of a protein family involved in
RT vesicular trafficking.";
RL J. Biol. Chem. 271:17183-17189(1996).
CC -1- FUNCTION: Could have a role in the budding of coatomer-coated and
CC other species of coated vesicles. Could bind cargo molecules to
CC collect them into budding vesicles.
CC -1- SUBUNIT: Interacts with ARFGAP1, inhibiting the GTPase-activating
CC activity of ARFGAP1 (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Golgi-derived
CC coatomer-coated vesicles.
CC -1- SIMILARITY: BELONGS TO THE EMP24/GP25L FAMILY.
CC -1- SIMILARITY: Contains 1 GOLD domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X92098; CAA63069.1; -.
DR GO; GO:0016021; C:Integral to membrane; TAS.
DR GO; GO:0005792; C:Microsome; TAS.
DR GO; GO:0005886; C:Plasma membrane; TAS.
DR GO; GO:0006886; P:intracellular protein transport; TAS.
DR InterPro; IPR000348; Emp24_gp25L_p24.
DR Pfam; PF01105; EMP24_GP25L; 1.
DR PROSITE; PS50866; GOLD; 1.
KM Transport; Protein transport; Transmembrane; Signal; Golgi stack.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 201 COP-COATED VESICLE MEMBRANE PROTEIN P24.
FT DOMAIN 21 168 LUMENAL (POTENTIAL).
FT TRANSMEM 169 189 POTENTIAL.
FT DOMAIN 190 201 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 30 112 GOLD.
SQ SEQUENCE 201 AA; 22761 MW; C452370E459DC894 CRC64;

Query Match 36.8%; Score 21; DB 1; Length 201;
Best Local Similarity 36.4%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXXGYYXF 29
| | | | |
DB 82 FAAHMDGTYKF 92

RESULT 9

P24_MOUSE STANDARD; PRT; 201 AA.
ID P24_MOUSE
AC Q9R0Q3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cop-coated vesicle membrane protein p24 precursor (p24A).
GN RNP24 OR SID394.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Seki N., Hattori A., Hayashi A., Kozuma S., Muramatsu M., Saito T.;
RT "Mouse mRNA for transmembrane protein."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Could have a role in the budding of coatomer-coated and
CC other species of coated vesicles. Could bind cargo molecules to
CC collect them into budding vesicles (By similarity).
CC -!- SUBUNIT: Interacts with ARFGAP1, inhibiting the GTPase-activating
CC activity of ARFGAP1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Golgi-derived
CC coatomer-coated vesicles (By similarity).
CC -!- SIMILARITY: BELONGS TO THE EMP24/GP25L FAMILY.
CC -!- SIMILARITY: Contains 1 GOLD domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
DR EMBL; AB025218; BAA84689.1; -.
DR MGD; MGI:1929269; Rnp24.
DR InterPro; IPR000348; Emp24_gp25L_p24.
DR Pfam; PF01105; EMP24_GP25L; 1.
DR PROSITE; PS50866; GOLD; 1.
KW Transport; Protein transport; Transmembrane; Signal; Golgi stack.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 1 201 COP-COATED VESICLE MEMBRANE PROTEIN P24.
FT DOMAIN 21 168 LUMENAL (POTENTIAL).
FT TRANSMEM 169 189 POTENTIAL.
FT DOMAIN 190 201 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 30 112 GOLD.
SQ SEQUENCE 201 AA; 22705 MW; F5DE259FC419CF85 CRC64;
Query Match 36.8%; Score 21; DB 1; Length 201;
Best Local Similarity 36.4%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
OY 19 FXXXXXXGYYXF 29
| | | | |
DB 82 FAAHMDGTYKF 92

RESULT 10

P24_RAT STANDARD; PRT; 201 AA.
ID P24_RAT
AC Q63524;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cop-coated vesicle membrane protein p24 precursor (p24A) (RNP21.4).
GN RNP24.

OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Pancreas;
RX MEDLINE=96291865; PubMed=8663407;
RA Blum R., Felck P., Puyte M., Vandekerckhove J., Klengel R.,
RA Nastainczyk W., Schulz I.;
RT "Timp21 and p24A, two type I proteins enriched in pancreatic
RT microsomal membranes, are members of a protein family involved in
RT vesicular trafficking."
RL J. Biol. Chem. 271:17183-17189(1996).
RN [2]

RP INTERACTION WITH ARFGAP1.
RX MEDLINE=21629435; PubMed=11748249;
RA Lanoix J., Owendijk J., Stark A., Szafer E., Cassel D., Dejgaard K.,
RA Weiss M., Nilsson T.;
RT "Sorting of Golgi resident proteins into different subpopulations of
RT COP1 vesicles: a role for ArfGAP1."
RL J. Cell Biol. 155:1199-1212(2001).
CC -!- FUNCTION: Could have a role in the budding of coatomer-coated and
CC other species of coated vesicles. Could bind cargo molecules to
CC collect them into budding vesicles.
CC -!- SUBUNIT: Interacts with ARFGAP1, inhibiting the GTPase-activating
CC activity of ARFGAP1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Golgi-derived
CC coatomer-coated vesicles.
CC -!- SIMILARITY: BELONGS TO THE EMP24/GP25L FAMILY.
CC -!- SIMILARITY: Contains 1 GOLD domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
DR EMBL; X92097; CAA63068.1; -.
DR InterPro; IPR000348; Emp24_gp25L_p24.
DR Pfam; PF01105; EMP24_GP25L; 1.
DR PROSITE; PS50866; GOLD; 1.
KW Transport; Protein transport; Transmembrane; Signal; Golgi stack.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 1 201 COP-COATED VESICLE MEMBRANE PROTEIN P24.
FT DOMAIN 21 168 LUMENAL (POTENTIAL).
FT TRANSMEM 169 189 POTENTIAL.
FT DOMAIN 190 201 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 30 112 GOLD.
SQ SEQUENCE 201 AA; 22733 MW; E35A3816429DDF9A CRC64;

Query Match 36.8%; Score 21; DB 1; Length 201;
Best Local Similarity 36.4%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXXGYYXF 29
| | | | |
DB 82 FAAHMDGTYKF 92

RESULT 11

YG26_HAEIN STANDARD; PRT; 238 AA.
ID YG26_HAEIN
AC P44278;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein HI1626.
GN HI1626.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

```
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA Mckenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudak D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -I- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -I- SIMILARITY: TO B.SUBTILIS YWIC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U32835; AAC23269.1; -.
DR PIR: I64038; I64038.
DR TIGR: H11626; -.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 19 39 POTENTIAL.
FT TRANSMEM 64 84 POTENTIAL.
FT TRANSMEM 85 105 POTENTIAL.
FT TRANSMEM 112 132 POTENTIAL.
FT TRANSMEM 141 161 POTENTIAL.
FT TRANSMEM 176 196 POTENTIAL.
FT TRANSMEM 218 238 POTENTIAL.
SQ SEQUENCE 238 AA; 27573 MW; 48B2034AB9F657A4 CRC64;

Query Match 36.8%; Score 21; DB 1; Length 238;
Best local Similarity 36.4%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXXGYXF 29
DB 122 FALAGMGXYF 132

RESULT 12
GLGC_AGRTS STANDARD; PRT; 420 AA.
AC Q8UBL5;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucose-1-phosphate adenylyltransferase (EC 2.7.7.27) (ADP-glucose
DE synthase) (ADP-glucose pyrophosphorylase) (ADPGlc Ppase).
GN GLGC OR ATU4076 OR AGR_L_1560.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kuttyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
```

```
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.F., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quorillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
CC -I- CATALYTIC ACTIVITY: ATP + alpha-D-glucose 1-phosphate =
CC diphosphate + ADP-glucose.
CC -I- PATHWAY: Glycogen biosynthesis; first step.
CC -I- SIMILARITY: BELONGS TO THE BACTERIAL AND PLANTS GLUCOSE-1-
CC PHOSPHATE ADENYLYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE009339; AAL44877.1; -.
DR EMBL: AE008276; AAK89353.1; -.
DR PIR: AG3057; AG3057.
DR PIR: G98228; G98228.
DR HAMAP: MF_00624; -.
DR InterPro: IPR005836; ADP_Glu_pyrop.
DR InterPro: IPR005835; NTP_transferase.
DR Pfam: PF00483; NTP_transferase; 1.
DR PROSITE: PS00808; ADP_GLC_PYROPHOSPH_1; 1.
DR PROSITE: PS00809; ADP_GLC_PYROPHOSPH_2; 1.
DR PROSITE: PS00810; ADP_GLC_PYROPHOSPH_3; 1.
KW Glycogen biosynthesis; Transferase; Nucleotidyltransferase;
KW Complete proteome.
SQ SEQUENCE 420 AA; 47043 MW; 6F6BC076EFC27484 CRC64;

Query Match 36.8%; Score 21; DB 1; Length 420;
Best local Similarity 36.4%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXXGYXF 29
DB 201 FALASMGIVF 211

RESULT 13
GLGC_AGRTU STANDARD; PRT; 420 AA.
ID GLGC_AGRTU
AC P39669;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucose-1-phosphate adenylyltransferase (EC 2.7.7.27) (ADP-glucose
DE synthase) (ADP-glucose pyrophosphorylase) (ADPGlc Ppase).
GN GLGC.
OS Agrobacterium tumefaciens;
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=358;
RN [1]
```


RP SEQUENCE FROM N.A.
RC STRAIN=A348;
RX MEDLINE=95047522; PubMed=7959036;
RA Urtaro A.D., Ugalde R.A.;
RT "A chromosomal cluster of genes encoding ADP-glucose synthetase,
RT glycogen synthase and phosphoglucomutase in Agrobacterium
RT tumefaciens.";
RL Gene 150:117-122(1994).
CC -1- CATALYTIC ACTIVITY: ATP + alpha-D-glucose 1-phosphate =
CC diphosphate + ADP-glucose.
CC -1- PATHWAY: Glycogen biosynthesis; first step.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL AND PLANTS GLUCOSE-1-
CC PHOSPHATE ADENYLYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF033856; AAD03473.1; -.
DR HAMAP; MF_00624; -; 1.
DR InterPro; IPR005836; ADP_Glu_pyrop.
DR InterPro; IPR005835; NTP_transferase.
DR Pfam; PF00483; NTP_transferase; 1.
DR PROSITE; PS00808; ADP_GLC_PYROPHOSPH_1; 1.
DR PROSITE; PS00809; ADP_GLC_PYROPHOSPH_2; 1.
DR PROSITE; PS00810; ADP_GLC_PYROPHOSPH_3; 1.
KW Glycogen biosynthesis; Transferase; Nucleotidyltransferase.
SQ SEQUENCE 420 AA; 47029 MW; F93F5EAB5996E698 CRC64;

QY Query Match 36.8%; Score 21; DB 1; Length 420;
Best Local Similarity 36.4%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 19 FXXXXXGXYYF 29
201 FALASMGIVVF 211

RESULT 14
GLGC_RHITR STANDARD; PRT; 420 AA.
AC O9EUT6;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucose-1-phosphate adenylyltransferase (EC 2.7.7.27) (ADP-glucose
DE synthase) (ADP-glucose pyrophosphorylase) (ADPGlc Ppase).
GN GLGC.
OS Rhizobium tropici.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=398;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CIAT899;
RX MEDLINE=21142510; PubMed=11208782;
RA Marroqui S., Zorreguieta A., Santamaria C., Temprano F., Soberon M.,
RA Megias M., Downie J.A.;
RT "Enhanced symbiotic performance by Rhizobium tropici glycogen synthase
RT mutants.";
RL J. Bacteriol. 183:854-864(2001).
CC -1- CATALYTIC ACTIVITY: ATP + alpha-D-glucose 1-phosphate =
CC diphosphate + ADP-glucose.
CC -1- PATHWAY: Glycogen biosynthesis; first step.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL AND PLANTS GLUCOSE-1-
CC PHOSPHATE ADENYLYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ291603; CAC17471.1; -.
DR HAMAP; MF_00624; -; 1.
DR InterPro; IPR005836; ADP_Glu_pyrop.
DR InterPro; IPR005835; NTP_transferase.
DR Pfam; PF00483; NTP_transferase; 1.
DR PROSITE; PS00808; ADP_GLC_PYROPHOSPH_1; 1.
DR PROSITE; PS00809; ADP_GLC_PYROPHOSPH_2; 1.
DR PROSITE; PS00810; ADP_GLC_PYROPHOSPH_3; 1.
KW Glycogen biosynthesis; Transferase; Nucleotidyltransferase.
SQ SEQUENCE 420 AA; 47006 MW; 4917D3C73EE08767 CRC64;

QY Query Match 36.8%; Score 21; DB 1; Length 420;
Best Local Similarity 36.4%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 19 FXXXXXGXYYF 29
201 FALASMGIVVF 211

RESULT 15
GLGC_RHILO STANDARD; PRT; 421 AA.
AC Q985F3; Q93QE7;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucose-1-phosphate adenylyltransferase (EC 2.7.7.27) (ADP-glucose
DE synthase) (ADP-glucose pyrophosphorylase) (ADPGlc Ppase).
GN GLGC OR MLR7588.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RA Lepek V.C., Tomatis P.E., Giambiagi S., Ugalde R.A.;
RT "Partial characterization of R.loti gene cluster for glycogen
RT metabolism: differences with A.tumefaciens.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
CC -1- CATALYTIC ACTIVITY: ATP + alpha-D-glucose 1-phosphate =
CC diphosphate + ADP-glucose.
CC -1- PATHWAY: Glycogen biosynthesis; first step.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL AND PLANTS GLUCOSE-1-
CC PHOSPHATE ADENYLYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF268969; AAK58595.1; -.

DR EMBL; AP003012; BAB54019.1; -.
DR HAMAP; MF_00624; -: 1.
DR InterPro; IPR005836; ADP_Glu_pyrop.
DR InterPro; IPR005835; NTP_transferase.
DR Pfam; PF00483; NTP_transferase; 1.
DR PROSITE; PS00808; ADP_GLC_PYROPHOSPH_1; 1.
DR PROSITE; PS00809; ADP_GLC_PYROPHOSPH_2; 1.
DR PROSITE; PS00810; ADP_GLC_PYROPHOSPH_3; 1.
KW Glycogen biosynthesis; Transferase; Nucleotidyltransferase;
KW Complete proteome.
FT CONFLICT 145 145 M -> L (IN REF. 1).
FT CONFLICT 154 154 N -> G (IN REF. 1).
FT CONFLICT 201 201 D -> E (IN REF. 1).
FT CONFLICT 227 227 E -> D (IN REF. 1).
FT CONFLICT 262 262 A -> M (IN REF. 1).
FT CONFLICT 287 287 I -> V (IN REF. 1).
FT CONFLICT 338 338 T -> S (IN REF. 1).
FT CONFLICT 355 355 T -> A (IN REF. 1).
FT CONFLICT 395 395 A -> V (IN REF. 1).
FT CONFLICT 411 411 I -> V (IN REF. 1).
SQ SEQUENCE 421 AA; 47036 MW; 6DD17FC327A876D9 CRC64;
Query Match 36.8%; Score 21; DB 1; Length 421;
Best Local Similarity 36.4%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 19 FXXXXXGXXF 29
Db 202 FALASMGIVVF 212

Search completed: January 12, 2004, 08:13:04
Job time : 4.11377 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:10 ; Search time 4.26946 Seconds
(without alignments)
1873.686 Million cell updates/sec

Title: US-10-085-167-3
Perfect score: 57
Sequence: 1 FXXXXXXXXXXXXXXXXXXXXXGXFXFX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 23 | 40.4 | 352 | 13 Q9DED3 | Q9ded3 oncorhynchu |
| 2 | 23 | 40.4 | 618 | 16 Q92LNG | Q92ln9 rhizobium m |
| 3 | 22 | 38.6 | 144 | 11 Q8BPU1 | Q8bpj1 mus musculu |
| 4 | 22 | 38.6 | 188 | 4 Q8NBUB | Q8nbu8 homo sapien |
| 5 | 22 | 38.6 | 224 | 4 Q8WU06 | Q8wuu6 homo sapien |
| 6 | 22 | 38.6 | 224 | 4 Q96K51 | Q96k51 homo sapien |
| 7 | 22 | 38.6 | 250 | 16 Q9KGP4 | Q9kqp4 bacillus ha |
| 8 | 22 | 38.6 | 290 | 16 Q8CSQ5 | Q8csq5 staphylococ |
| 9 | 22 | 38.6 | 300 | 16 Q9CNG6 | Q9cng6 pasteurella |
| 10 | 22 | 38.6 | 302 | 16 Q8DBA0 | Q8dba0 vibrio vuln |
| 11 | 22 | 38.6 | 302 | 16 Q9KQ59 | Q9kq59 vibrio chol |
| 12 | 22 | 38.6 | 323 | 17 Q8PXN5 | Q8pxn5 methanosarc |
| 13 | 22 | 38.6 | 325 | 17 Q8PSM8 | Q8psm8 methanosarc |
| 14 | 22 | 38.6 | 355 | 5 Q966G5 | Q966g5 caenorhabdi |
| 15 | 22 | 38.6 | 384 | 16 Q8EJZ4 | Q8ejz4 shewanella |
| 16 | 22 | 38.6 | 404 | 17 Q8U1U7 | Q8u1u7 pyrococcus |

| | | | | | |
|----|----|------|------|-----------|--------------------|
| 17 | 22 | 38.6 | 409 | 11 Q91ZK1 | Q91zk1 ratius norv |
| 18 | 22 | 38.6 | 421 | 17 Q97Y43 | Q97y43 sulfolobus |
| 19 | 22 | 38.6 | 423 | 17 Q97UD8 | Q97ud8 sulfolobus |
| 20 | 22 | 38.6 | 440 | 16 Q8XPA4 | Q8xpa4 clostridium |
| 21 | 22 | 38.6 | 441 | 5 Q18017 | Q18017 caenorhabdi |
| 22 | 22 | 38.6 | 529 | 10 Q9FGF7 | Q9fgf7 arabidopsis |
| 23 | 22 | 38.6 | 679 | 16 Q8XW21 | Q8xw21 ralstonia s |
| 24 | 22 | 38.6 | 787 | 5 Q9GSD4 | Q9gsd4 plasmodium |
| 25 | 22 | 38.6 | 868 | 5 Q6Q984 | Q6q984 dictyosteli |
| 26 | 22 | 38.6 | 1127 | 2 Q8L3A4 | Q8l3a4 bacillus st |
| 27 | 22 | 38.6 | 1457 | 2 Q9F289 | Q9f289 yersinia pe |
| 28 | 22 | 38.6 | 1677 | 10 Q9AS94 | Q9as94 oryza sativ |
| 29 | 22 | 38.6 | 1913 | 5 Q9GR96 | Q9gr96 leucophaea |
| 30 | 22 | 38.6 | 3705 | 2 Q9F285 | Q9f285 yersinia pe |
| 31 | 22 | 38.6 | 3705 | 16 Q8ZHA1 | Q8zha1 yersinia pe |
| 32 | 22 | 38.6 | 3710 | 16 Q8CZU2 | Q8czu2 yersinia pe |
| 33 | 22 | 38.6 | 4976 | 2 Q87314 | Q87314 mycobacteri |
| 34 | 22 | 38.6 | 5020 | 16 Q8E9W3 | Q8e9w3 shewanella |
| 35 | 21 | 36.8 | 45 | 16 Q8F630 | Q8f630 leptospira |
| 36 | 21 | 36.8 | 109 | 5 Q9NHK7 | Q9nhk7 plasmodium |
| 37 | 21 | 36.8 | 109 | 5 Q9NHL5 | Q9nhl5 plasmodium |
| 38 | 21 | 36.8 | 113 | 5 Q8NMK3 | Q8nmk3 dictyosteli |
| 39 | 21 | 36.8 | 135 | 16 Q8FAA7 | Q8faa7 escherichia |
| 40 | 21 | 36.8 | 154 | 2 Q9L5R0 | Q9l5r0 salmonella |
| 41 | 21 | 36.8 | 154 | 16 Q935R2 | Q935r2 salmonella |
| 42 | 21 | 36.8 | 156 | 16 Q9Z7P4 | Q9z7p4 chlamydia p |
| 43 | 21 | 36.8 | 159 | 11 Q8C1Z9 | Q8c1z9 mus musculu |
| 44 | 21 | 36.8 | 169 | 16 Q8E5W2 | Q8e5w2 streptococc |
| 45 | 21 | 36.8 | 169 | 16 Q8E085 | Q8e085 streptococc |

ALIGNMENTS

RESULT 1

Q9DED3 PRELIMINARY; PRT; 352 AA.

Q9DED3

AC Q9DED3;

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)

DE Inhibin.

GN INH.

OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;

OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

OX NCBI_TaxID=8022;

RN [1]

RP SEQUENCE FROM N.A.

RA Tada T., Endo M., Hirono I., Takashima F., Aoki T.;

RT "Differential expression and cellular localization of activin and

RT inhibin mRNA in the rainbow trout ovary and testis.";

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.

DR EMBL; AB044566; BAB19272.1; -.

DR InterPro; IPR002405; Inhibin_alpha.

DR InterPro; IPR001839; TGFb.

DR Pfam; PF00019; TGF-beta; 1.

DR PRINTS; PRO0669; INHIBINA.

DR ProDom; PD000357; TGFb; 1.

DR SMART; SM00204; TGFb; 1.

DR PROSITE; PS00250; TGF_BETA_1; 1.

SQ SEQUENCE 352 AA; 39711 MW; ED9CCCEC860F912ED CRC64;

Query Match 40.4%; Score 23; DB 13; Length 352;
Best Local Similarity 36.4%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXXXXFX 29

DB 326 FTTTSDGGYSF 336

```
RESULT 2
Q92LN9          PRELIMINARY;      PRT;      618 AA.
ID Q92LN9
AC Q92LN9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein R02996.
GN R02996 OR SMC03097.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetlelle D., Puhler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RA "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL; AL591792; CAC47575.1; -.
DR InterPro; IPR00184; Bac_surfAg_D15.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR001412; tRNA-synt_1.
DR Pfam; PF01103; Bac_surface_Ag; 1.
DR PROSITE; PS00178; AA tRNA_LIGASE_I; 1.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
DR KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 618 AA; 65656 MW; E51276BB29B1FB78 CRC64;

Query Match          40.4%; Score 23; DB 16; Length 618;
Best Local Similarity 36.4%; Pred. No. 9.1e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGYXF 29
DB 464 FEASASGYAF 474

RESULT 3
Q8BPJ1          PRELIMINARY;      PRT;      144 AA.
ID Q8BPJ1
AC Q8BPJ1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Runt related transcription factor 3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Eye;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK053910; BAC35587.1; -.
SQ SEQUENCE 144 AA; 14731 MW; EC4789698ADBD9D7 CRC64;

Query Match          38.6%; Score 22; DB 11; Length 144;
Best Local Similarity 36.4%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```
QY 19 FXXXXXXGYXF 29
DB 66 FYGASSGSYGF 76

RESULT 4
Q8NBUS          PRELIMINARY;      PRT;      188 AA.
ID Q8NBUS
AC Q8NBUS;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ90737.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
RA Masuko Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RT "NEDO human cDNA sequencing project."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK075218; BAC11479.1; -.
DR InterPro; IPR000348; Emp24_gp25L_p24.
DR Pfam; PF01105; EMP24_GP25L; 1.
KW Hypothetical protein.
SQ SEQUENCE 188 AA; 21233 MW; 15A10CC4BD553DB3 CRC64;

Query Match          38.6%; Score 22; DB 4; Length 188;
Best Local Similarity 36.4%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGYXF 29
DB 98 FTASKNGTYKF 108

RESULT 5
Q8WU6           PRELIMINARY;      PRT;      224 AA.
ID Q8WU6
AC Q8WU6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ90481.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Isegai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
RA Masuko Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RT "NEDO human cDNA sequencing project."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC019349; AAH19349.1; -.
DR EMBL; AK074962; BAC11318.1; -.
DR InterPro; IPR000348; Emp24_gp25L_p24.
DR Pfam; PF01105; EMP24_GP25L; 1.
KW Hypothetical protein.
```

SEQ SEQUENCE 224 AA; 25171 MW; 9D0D2110579837B3 CRC64;

Query Match 38.6%; Score 22; DB 4; Length 224;
Best Local Similarity 36.4%; Pred. No. 5.8e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29
| | | | |
Db 98 FTASKNGTYKF 108

RESULT 6

Q96K51 PRELIMINARY; PRT; 224 AA.

AC Q96K51;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ14606.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Magatsuma M., Hosolri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK027512; BAB55166.1; -
DR InterPro; IPR000348; Emp24_gp25L_p24.
DR Pfam; PF01105; EMP24_GP25L.1.
KW Hypothetical protein.
SQ SEQUENCE 224 AA; 25181 MW; 9D0D2113989837B3 CRC64;

Query Match 38.6%; Score 22; DB 4; Length 224;
Best Local Similarity 36.4%; Pred. No. 5.8e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29
| | | | |
Db 98 FTASKNGTYKF 108

RESULT 7

Q9KGP4 PRELIMINARY; PRT; 250 AA.

AC Q9KGP4;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein BH0013.
GN BH0013.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis."
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001507; BAB03732.1; -
DR InterPro; IPR006520; Tail_N.

DR TIGRFAMS; TIGR01633; phi3626_gp14_N; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 250 AA; 28481 MW; DCD4B472571070FB CRC64;

Query Match 38.6%; Score 22; DB 16; Length 250;
Best Local Similarity 36.4%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29
| | | | |
Db 132 FTFSGAGSYDF 142

RESULT 8

O8CSQ5 PRELIMINARY; PRT; 290 AA.

AC O8CSQ5;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Conserved hypothetical protein.
GN SE1006.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 12228;
RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,
RA Chen Z., Wen Y.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016747; AA004603.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 290 AA; 33745 MW; A13E6A424943B195 CRC64;

Query Match 38.6%; Score 22; DB 16; Length 290;
Best Local Similarity 36.4%; Pred. No. 7.6e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29
| | | | |
Db 171 FKTRSGVYTF 181

RESULT 9

Q9CNH6 PRELIMINARY; PRT; 296 AA.

AC Q9CNH6;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein PM0455.
GN PM0455.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genome sequence of Pasteurella multocida Pm70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; AE006081; AAK02539.1; -
DR InterPro; IPR006314; Dyp_peroxidase.
DR TIGRFAMS; TIGR01413; Dyp_perox_fam; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 296 AA; 32951 MW; C4B616FAB4E40197 CRC64;

Query Match 38.6%; Score 22; DB 16; Length 296;
Best Local Similarity 36.4%; Pred. No. 7.8e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29
| | |
Db 276 FTKAVTGAYFF 286

RESULT 10

Q8DBA0 PRELIMINARY; PRT; 300 AA.

AC Q8DBA0;

DT 01-MAR-2003 (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE Tyra protein.

GN VV11922.

OS Vibrio vulnificus.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;

OC Vibrionaceae; Vibrio.

OX NCBI_TaxID=672;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CMCP6;

RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,

RT Choy H.E.;

RL "Complete genome sequence of Vibrio vulnificus CMCP6.";

DR Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.

KW EMBL; AE016803; AA010323.1; -.

KM Complete proteome.

SQ SEQUENCE 300 AA; 33262 MW; 7369A026F23B6298 CRC64;

QY 19 FXXXXXXGXXF 29
| | |
Db 277 FTKAVTGAYFF 287

Query Match

Best Local Similarity 38.6%; Score 22; DB 16; Length 300;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

RESULT 11
Q9KQ59 PRELIMINARY; PRT; 302 AA.

AC Q9KQ59;

DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE Tyra protein.

GN VC2145.

OS Vibrio cholerae.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;

OC Vibrionaceae; Vibrio.

OX NCBI_TaxID=666;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=El Tor N16961 / Serotype O1;

RX MEDLINE=20406833; PubMed=10952301;

RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,

RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,

RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,

RA Ermolaeva M.D., Vamathevan J., Bess S., Qin H., Dragoi I., Sellers P.,

RA McDonald L., Utterback T., Fleischmann R.D., Nieman W.C., White O.,

RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,

RA Fraser C.M.;

RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio

RT cholerae.";

RL Nature 406:477-483(2000).

DR EMBL; AE004287; AAF95290.1; -.

DR TIGR; VC2145; -.

DR InterPro; IPR006314; DYP_peroxidase.

DR TIGRFAMs; TIGR01413; DYP_perox_fam; 1.

KM Complete proteome.

SQ SEQUENCE 302 AA; 33496 MW; B45BAC9388BF2F7E CRC64;

Query Match 38.6%; Score 22; DB 16; Length 302;
Best Local Similarity 36.4%; Pred. No. 7.9e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29
| | |
Db 277 FTKAVTGAYFF 287

RESULT 12

Q8PXN5 PRELIMINARY; PRT; 323 AA.

AC Q8PXN5;

DT 01-OCT-2002 (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE Conserved protein.

GN M1181.

OS Methanosarcina mazel (Methanosarcina frisia).

OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;

OC Methanosarcinaceae; Methanosarcina.

OX NCBI_TaxID=2209;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Goel / Go1 / ATCC BAA-199 / DSM 3647 / OCM 88;

RX MEDLINE=22120827; PubMed=12125824;

RA Deppenmeier U., Johann A., Hartesch T., Merkl R., Schmitz R.A.,

RA Martinez-Arias R., Henne A., Wierzer A., Baumer S., Jacobi C.,

RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,

RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,

RA Fritz H.-J., Gottschalk G.;

RT "The genome of Methanosarcina mazel: evidence for lateral gene

RT transfer between Bacteria and Archaea.";

RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).

DR EMBL; AE013348; AAM30877.1; -.

KW Complete proteome.

SQ SEQUENCE 323 AA; 36826 MW; 8DDFE82C44BB87AE CRC64;

Query Match

Best Local Similarity 38.6%; Score 22; DB 17; Length 323;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29
| | |
Db 178 FAYYLSGAYAF 188

RESULT 13

Q8PSM8 PRELIMINARY; PRT; 325 AA.

AC Q8PSM8;

DT 01-OCT-2002 (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE Hypothetical protein MM3051.

GN Methanosarcina mazel (Methanosarcina frisia).

OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;

OC Methanosarcinaceae; Methanosarcina.

OX NCBI_TaxID=2209;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Goel / Go1 / ATCC BAA-199 / DSM 3647 / OCM 88;

RX MEDLINE=22120827; PubMed=12125824;

RA Deppenmeier U., Johann A., Hartesch T., Merkl R., Schmitz R.A.,

RA Martinez-Arias R., Henne A., Wierzer A., Baumer S., Jacobi C.,

RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,

RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,

RA Fritz H.-J., Gottschalk G.;

RT "The genome of Methanosarcina mazel: evidence for lateral gene

RT transfer between Bacteria and Archaea.";

RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).

DR EMBL; AE013560; AAM32747.1; -.

KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 325 AA; 36804 MW; 4BA35BD04B1D61CA CRC64;
Query Match 38.6%; Score 22; DB 17; Length 325;
Best Local Similarity 36.4%; Pred. No. 8.5e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 19 FXXXXXXGXYXF 29
Db 178 FAYYLSGAYAF 188
RESULT 14
Q966G5 PRELIMINARY; PRT; 355 AA.
AC Q966G5; 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein R13D7.10.
GN R13D7.10.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Murray J., Wohlmann P., Elliott G.;
RT "The sequence of C. elegans cosmid R13D7.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006680; AAK72298.1; -.
DR WormPep; R13D7.10; CE18155.
DR InterPro; IPR002651; DUF32.
DR Pfam; PF01748; DUF32; 1.
SQ SEQUENCE 355 AA; 40110 MW; DF4CF926589473BF CRC64;
Query Match 38.6%; Score 22; DB 5; Length 355;
Best Local Similarity 36.4%; Pred. No. 9.3e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 19 FXXXXXXGXYXF 29
Db 251 FHATVAGLYSF 261
RESULT 15
Q8EJZ4 PRELIMINARY; PRT; 384 AA.
AC Q8EJZ4; 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Outer membrane porin, putative.
GN SO0312.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Nealon K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AE015479; AAN53397.1; -.
DR TIGR; SO0312; -.
KW Porin; Complete proteome.
SQ SEQUENCE 384 AA; 42311 MW; 81E68F1A06C15B84 CRC64;
Query Match 38.6%; Score 22; DB 16; Length 384;
Best Local Similarity 36.4%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 19 FXXXXXXGXYXF 29
Db 144 FAGYSVGAYAF 154

Search completed: January 12, 2004, 08:19:19
Job time : 9.26946 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:09 ; Search time 4.67166 Seconds
(without alignments)
1053.272 Million cell updates/sec

Title: US-10-085-167-3
Perfect score: 57
Sequence: 1 FXXXXXXXXXXXXXXXXXXXXGXXFFX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

1: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA2002.DAT:*
24: /SIDS1/gcgcdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|--------------------|
| 1 | 22 | 38.6 | 163 | 19 | AAW70458 | Human T1-receptor |
| 2 | 22 | 38.6 | 188 | 22 | AAW93624 | Human polypeptide, |
| 3 | 22 | 38.6 | 215 | 19 | AAW70459 | Human T1-receptor |
| 4 | 22 | 38.6 | 215 | 23 | ABP65162 | Hypoxia-regulated |
| 5 | 22 | 38.6 | 219 | 22 | ABB11780 | Human T1 receptor |
| 6 | 22 | 38.6 | 224 | 22 | AAW93844 | Human polypeptide, |
| 7 | 22 | 38.6 | 224 | 22 | AAW92820 | Human protein sequ |
| 8 | 22 | 38.6 | 231 | 24 | ABR41674 | Human DITHP cell m |
| 9 | 22 | 38.6 | 266 | 22 | AAW82486 | S. epidermidis ope |

| | | | | | | |
|----|----|------|-----|----|----------|--------------------|
| 10 | 22 | 38.6 | 292 | 23 | ABP39297 | Staphylococcus epi |
| 11 | 22 | 38.6 | 540 | 22 | AAU57054 | Propionibacterium |
| 12 | 21 | 36.8 | 20 | 22 | AAW73384 | T. harzianum beta- |
| 13 | 21 | 36.8 | 111 | 20 | AAW89961 | Antigen from clust |
| 14 | 21 | 36.8 | 128 | 23 | ABG76563 | HCV E1 antigen mon |
| 15 | 21 | 36.8 | 151 | 20 | AAV37071 | Protein which is s |
| 16 | 21 | 36.8 | 154 | 24 | ABP75824 | Human secretory po |
| 17 | 21 | 36.8 | 156 | 20 | AAV35280 | Chlamydia pneumoni |
| 18 | 21 | 36.8 | 169 | 23 | ABP27601 | Streptococcus poly |
| 19 | 21 | 36.8 | 183 | 23 | ABP69421 | Human polypeptide |
| 20 | 21 | 36.8 | 201 | 22 | AAW74750 | Human secreted pro |
| 21 | 21 | 36.8 | 201 | 22 | AAW74750 | Human secreted pro |
| 22 | 21 | 36.8 | 201 | 22 | AAW74760 | Human secreted pro |
| 23 | 21 | 36.8 | 201 | 23 | ABG65324 | Human albumin fusi |
| 24 | 21 | 36.8 | 201 | 23 | ABG65326 | Human albumin fusi |
| 25 | 21 | 36.8 | 201 | 23 | ABG65327 | Human albumin fusi |
| 26 | 21 | 36.8 | 208 | 23 | ABP69422 | Human polypeptide |
| 27 | 21 | 36.8 | 210 | 18 | AAW20426 | H. pylori surface |
| 28 | 21 | 36.8 | 215 | 18 | AAW20968 | H. pylori transmem |
| 29 | 21 | 36.8 | 254 | 18 | AAU31151 | Novel human secret |
| 30 | 21 | 36.8 | 271 | 21 | AAW76114 | Arabidopsis thalia |
| 31 | 21 | 36.8 | 271 | 21 | AAW76114 | Arabidopsis thalia |
| 32 | 21 | 36.8 | 277 | 21 | AAW76113 | Arabidopsis thalia |
| 33 | 21 | 36.8 | 277 | 21 | AAW76113 | Arabidopsis thalia |
| 34 | 21 | 36.8 | 280 | 23 | ABP29735 | Streptococcus poly |
| 35 | 21 | 36.8 | 281 | 22 | ABG23769 | Novel human diagno |
| 36 | 21 | 36.8 | 295 | 21 | AAW51316 | Arabidopsis thalia |
| 37 | 21 | 36.8 | 296 | 21 | AAW07612 | Arabidopsis thalia |
| 38 | 21 | 36.8 | 300 | 21 | AAV91317 | Group B Streptococ |
| 39 | 21 | 36.8 | 300 | 23 | ABP26518 | Streptococcus poly |
| 40 | 21 | 36.8 | 307 | 21 | AAW18332 | Plasmodium falcipa |
| 41 | 21 | 36.8 | 336 | 23 | ABP62884 | Human polypeptide |
| 42 | 21 | 36.8 | 382 | 22 | ABW60990 | Drosophila melanog |
| 43 | 21 | 36.8 | 408 | 22 | AAW79473 | Corynebacterium gl |
| 44 | 21 | 36.8 | 441 | 22 | AAW92180 | C glutamicum prote |
| 45 | 21 | 36.8 | 473 | 23 | ABW98126 | Human PPMV incyte |

ALIGNMENTS

| | | | |
|----------|---|---|--|
| RESULT 1 | | AAW70458 standard; Protein; 163 AA. | |
| AAW70458 | 10-DEC-1998 (first entry) | | |
| XX | Human T1-receptor ligand III splice variant 1. | | |
| DE | Human T1-receptor ligand III; T1-R ligand III; atherosclerosis; | | |
| XX | Human T1-receptor ligand III; T1-R ligand III; atherosclerosis; | | |
| KW | autoimmune disease; inflammation; metabolic dysfunction; | | |
| KW | immune-regulated disorder. | | |
| XX | | | |
| OS | Homo sapiens. | | |
| XX | | | |
| FH | Key | Location/Qualifiers | |
| FT | Peptide | 1..24 | |
| FT | | /note= "Signal peptide" | |
| FT | Protein | 25..163 | |
| XX | | /note= "T1-R ligand III splice variant 1" | |
| PN | WO9838311-A1. | | |
| XX | | | |
| PD | 03-SEP-1998. | | |
| XX | | | |
| PF | 26-FEB-1998; | 98WO-US03483. | |
| XX | | | |
| PR | 28-FEB-1997; | 97US-0039483. | |
| XX | | | |
| PA | (HUMA-) HUMAN GENOME SCI INC. | | |
| XX | | | |

PI Gentz RL, Ni J, Ruben SM;
XX WPI; 1998-495398/42.
DR N-PSDB; AAV33461.
XX
PT New human T1-receptor ligand III splice variant(s) - useful for,
PT e.g. detecting and treating immune system related disorders such as
PT cancer and inflammation
XX
PS Claim 1; Fig 1; 115pp; English.
XX
CC The invention provides novel human T1-receptor ligand III (T1-R
CC ligand III) splice variants and antibodies raised against these
CC proteins. The present sequence represents the T1-R ligand III splice
CC variant 1 protein sequence. The T1-R ligand III splice variants are
CC claimed useful for screening agonists and antagonists. They are also
CC claimed useful for treating disorders such as atherosclerosis,
CC autoimmune disease, inflammation, metabolic dysfunction and
CC immune-regulated disorders.
XX
SQ Sequence 163 AA;

Query Match 38.6%; Score 22; DB 19; Length 163;
Best Local Similarity 36.4%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXXF 29
| | | | |
DB 89 FTASKNGTYKF 99

RESULT 2
AAM93624
ID AAM93624 standard; Protein; 188 AA.
XX
AC AAM93624;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide, SEQ ID NO: 3460.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
DR N-PSDB; AAK94557.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
PS Claim 8; SEQ ID NO 3460; 1380pp + sequence listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by

CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a polypeptide
CC encoded by a full length human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 188 AA;

Query Match 38.6%; Score 22; DB 22; Length 188;
Best Local Similarity 36.4%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXXF 29
| | | | |
DB 98 FTASKNGTYKF 108

RESULT 3
AAW70459
ID AAW70459 standard; Protein; 215 AA.
XX
AC AAW70459;
XX
DT 10-DEC-1998 (first entry)
XX
DE Human T1-receptor ligand III splice variant 2.
XX
KW Human T1-receptor ligand III; T1-R ligand III; atherosclerosis;
KW autoimmune disease; inflammation; metabolic dysfunction;
KW immune-regulated disorder.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..24
FT Protein /note= "Signal peptide"
FT /note= "T1-R ligand III splice variant 2"
XX
PN WO9838311-A1.
XX
PD 03-SEP-1998.
XX
PF 26-FEB-1998; 98WO-US03483.
XX
PR 28-FEB-1997; 97US-0039483.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Gentz RL, Ni J, Ruben SM;
XX
DR WPI; 1998-495398/42.
DR N-PSDB; AAV33462.
XX
PT New human T1-receptor ligand III splice variant(s) - useful for,
PT e.g. detecting and treating immune system related disorders such as
PT cancer and inflammation
XX
PS Claim 1; Fig 2; 115pp; English.
XX
CC The invention provides novel human T1-receptor ligand III (T1-R
CC ligand III) splice variants and antibodies raised against these
CC proteins. The present sequence represents the T1-R ligand III splice
CC variant 2 protein sequence. The T1-R ligand III splice variants are
CC claimed useful for screening agonists and antagonists. They are also
CC claimed useful for treating disorders such as atherosclerosis,
CC autoimmune disease, inflammation, metabolic dysfunction and
CC immune-regulated disorders.
XX
SQ Sequence 215 AA;

Query Match 38.6%; Score 22; DB 19; Length 215;
Best Local Similarity 36.4%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXGXXF 29
| | | |
Db 89 FTASKNGTYKF 99

RESULT 4

ABP65162 ID ABP65162 standard; Protein; 215 AA.

XX AC ABP65162;

XX DT 12-NOV-2002 (first entry)

XX DE Hypoxia-regulated protein #36.

XX KW Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;

KW antiinflammatory; vulnerary; gynecological; ophthalmological; vaccine;

KW hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;

KW ischaemic condition; reperfusion injury; retinopathy; neonatal stress;

KW preeclampsia; atherosclerosis; inflammatory condition; wound healing;

KW inflammation; erythropoiesis; hair loss; human.

XX OS Homo sapiens.

XX PN WO200246465-A2.

XX PD 13-JUN-2002.

XX PF 10-DEC-2001; 2001WO-GB05458.

XX PR 08-DEC-2000; 2000GB-0030076.

XX PR 08-FEB-2001; 2001GB-0003156.

XX PR 25-OCT-2001; 2001GB-0025666.

XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.

XX PI White J, Mundy CR, Ward NR, Krige D, Kingman SM, Harris RA;

XX PI Rayner WN;

XX DR WPI; 2002-627238/67.

XX PT Identifying a gene involved in disease for treating hypoxia-regulated

XX PT conditions, comprises comparing the transcriptome/proteome of two cell

XX PT types under different conditions and identifying a differentially

XX PT regulated gene -

XX PS Claim 35; Page 355; 538pp; English.

XX CC The present invention relates to methods for identifying genes and

XX CC proteins that are implicated in a specific disease or physiological

XX CC condition. The method comprises comparing the transcriptome/proteome of a

XX CC specialised cell type implicated in a disease or condition with that of a

XX CC second specialised cell type, under two experimental conditions, and

XX CC identifying a gene that is differentially regulated in the two

XX CC specialised cell types under experimental conditions. ABV77873-ABV78116

XX CC and ABP65061-ABP65257 were identified using the methods of the invention.

Best Local Similarity 36.4%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
OY 19 FXXXXXGXXF 29
| | | |
Db 90 FTASKNGTYKF 100

RESULT 5

ABB11780 ID ABB11780 standard; peptide; 219 AA.

XX AC ABB11780;

XX DT 11-JAN-2002 (first entry)

XX DE Human T1 receptor ligand III homologue, SEQ ID NO:2150.

XX KW Human; cytokine; cell proliferation; cell differentiation; growth factor;

KW haematopoiesis regulation; tissue growth; immunomodulator; activin;

KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;

KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;

KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;

KW chronic inflammatory condition; proliferative retinopathy;

KW atherosclerosis; coronary heart disease; arterial ischaemia;

KW bone disorder; osteoporosis; vascular growth disorder;

KW tissue regeneration; wound healing; infection; immune disorder;

KW cell culture; drug screening; gene therapy; antiinflammatory;

KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;

KW cyostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;

KW antifungal; vulnerary; antilulcer.

XX OS Homo sapiens.

XX PN WO200157188-A2.

XX PD 09-AUG-2001.

XX PF 05-FEB-2001; 2001WO-US03800.

XX PR 03-FEB-2000; 2000US-0496914.

XX PR 27-APR-2000; 2000US-0560875.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX PI WPI; 2001-457740/49.

XX DR N-PSDB; ABA09024.

XX PT Human proteins and DNA encoding sequences useful for preventing,

XX PT treating or ameliorating a medical condition in a mammalian subject

XX PT e.g. arthritis and cancer -

XX PS Claim 20; Page 245; 1963pp; English.

XX CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and

XX CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The

XX CC invention also relates to vectors and recombinant host cells comprising a

XX CC nucleotide of the invention, methods of producing the novel polypeptides,

XX CC antibodies against the polypeptides, methods of detecting the nucleotides

XX CC or polypeptides in a sample, and methods of identifying compounds which

XX CC bind to polypeptides of the invention. Although novel, many of the

XX CC polypeptides of the invention have homology to known proteins, thereby

XX CC giving an insight into their probable biological activities, and hence

XX CC potential therapeutic applications. The polypeptides of the invention may

XX CC have various activities, including cytokine, cell proliferation or cell

XX CC differentiation activities; stem cell growth factor activity;

XX CC haematopoiesis regulatory activity; tissue growth activity;

XX CC immunomodulatory activity; activin- or inhibin-related activities;

XX CC chemotactic or chemokinetic activities; haemostatic, thrombotic or

XX CC thrombolytic activities; receptor or ligand activities; or may be

XX CC involved in oncogenesis, cancer cell proliferation or metastasis.

Query Match 38.6%; Score 22; DB 23; Length 215;

CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.

XX SQ Sequence 219 AA;

Query Match 38.6%; Score 22; DB 22; Length 219;

Best Local Similarity 36.4%; Pred. No. 2.8e+02;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXXGXXF 29

Db 93 FTASKNGTYKF 103

RESULT 6

AAM93844

ID AAM93844 standard; Protein; 224 AA.

XX AC AAM93844;

XX DT 06-NOV-2001 (first entry)

XX DE Human polypeptide, SEQ ID NO: 3923.

XX KW Human; full length cDNA; cDNA synthesis; oligo-capping.

XX OS Homo sapiens.

XX PN EP1130094-A2.

XX PD 05-SEP-2001.

XX PF 07-JUL-2000; 2000EP-0114089.

XX PR 08-JUL-1999; 99JP-0194486.

XX PR 11-JAN-2000; 2000JP-0118774.

XX PR 02-MAY-2000; 2000JP-0183765.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX DR N-PSDB; AAK94800.

XX PT 830 Primers useful for synthesizing full length cDNA clones and their

XX PS use in genetic manipulation -

XX PS Claim 8; SEQ ID NO 3923; 1380pp + sequence listing; English.

XX CC The invention relates to primers for synthesising full length cDNA

XX CC clones. 830 cDNA molecules encoding a human protein have been

XX CC isolated, and nucleotide sequences of 5'- and 3'-ends of the cDNA

XX CC molecules have been determined. Primers for synthesising the full length

CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a polypeptide
CC encoded by a full length human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.

XX SQ Sequence 224 AA;

Query Match 38.6%; Score 22; DB 22; Length 224;

Best Local Similarity 36.4%; Pred. No. 2.9e+02;

Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXXGXXF 29

Db 98 FTASKNGTYKF 108

RESULT 7

AAB92820

ID AAB92820 standard; Protein; 224 AA.

XX AC AAB92820;

XX DT 26-JUN-2001 (first entry)

XX DE Human protein sequence SEQ ID NO:11348.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 99JP-0248036.

XX PR 27-AUG-1999; 99JP-0300253.

XX PR 11-JAN-2000; 2000JP-0118776.

XX PR 02-MAY-2000; 2000JP-0183767.

XX PR 09-JUN-2000; 2000JP-0241899.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX DR WPI; 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602

XX PT full-length cDNAs defined in the specification, and for the detection

XX PT and/or diagnosis of the abnormality of the proteins encoded by the

XX PT full-length cDNAs -

XX PS Claim 8; SEQ ID 11348; 2537pp + CD ROM; English.

XX CC The present invention describes primer sets for synthesising 5602

XX CC full-length cDNAs defined in the specification. Where a primer set

XX CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

XX CC to the complementary strand of a polynucleotide which comprises one of

XX CC the 5602 nucleotide sequences defined in the specification, where the

XX CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

XX CC of an oligonucleotide comprising a sequence complementary to the

CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 224 AA;

Query Match 38.6%; Score 22; DB 22; Length 224;
Best Local Similarity 36.4%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXVXF 29
| | | | |
Db 98 FTASKNGTYKF 108

RESULT 8
ABR41674 ID ABR41674 standard; Protein; 231 AA.
XX AC ABR41674;
XX DT 02-JUN-2003 (first entry)
XX DE Human DITHP cell membrane protein.
XX
KW Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
KW cancer; cell proliferative disorder; autoimmune disorder;
KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;
KW neurological disorder; gastrointestinal disorder; transport disorder;
KW connective tissue disorder; drug screening; proteome analysis;
KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
KW disease model; toxicological testing; transcript imaging;
KW cell membrane protein.
XX
OS Homo sapiens.
XX
PN WO200297031-A2.
XX
PD 05-DEC-2002.
XX
PF 27-MAR-2002; 2002WO-US10056.
XX
PR 28-MAR-2001; 2001US-279619P.
PR 29-MAR-2001; 2001US-280067P.
PR 29-MAR-2001; 2001US-280068P.
PR 16-MAY-2001; 2001US-291280P.
PR 17-MAY-2001; 2001US-291829P.
PR 17-MAY-2001; 2001US-291849P.
PR 19-JUN-2001; 2001US-299428P.
PR 20-JUN-2001; 2001US-299776P.
PR 20-JUN-2001; 2001US-300001P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Dafo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
PI Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstein EH;
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI Flores V, Marwaha R, Lo A, Ian RY, Urashka ME;
XX
DR WPI; 2003-129518/12.
DR N-PSDB; ACC46611.
XX
PT Novel human diagnostic and therapeutic polypeptide useful for
PT identifying test compound which specifically binds to a polypeptide
PT encoded by human diagnostic and therapeutic polynucleotide, and to
PT induce antibodies

XX Claim 27; SEQ ID No 1209; 591bp; English.
PS
XX The invention relates to novel human diagnostic and therapeutic
CC polynucleotides designated dithp (ACC46080-ACC46749) and to their
CC encoded proteins (DITHP; ABR41136-ABR41812). The invention also relates
CC to polynucleotide sequences at least 90% identical to the dithp cDNA
CC sequences of the invention; recombinant vectors, host cells and
CC transgenic organisms comprising a dithp nucleic acid sequence; the
CC recombinant production of DITHP proteins; antibodies specific for DITHP
CC proteins; microarrays comprising dithp nucleic acid sequences; methods
CC of detecting dithp nucleotide and protein sequences; methods of screening
CC for compounds which specifically bind a DITHP protein; and methods of
CC assessing the toxicity of test compounds using a dithp hybridisation
CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the
CC diagnosis of a wide variety of conditions including cancer and other cell
CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,
CC viral, fungal or parasitic infections; hormonal disorders; metabolic
CC disorders; neurological disorders; gastrointestinal disorders; transport
CC disorders; and connective tissue disorders. They may also be used to
CC screen for modulators of protein activity or gene expression. DITHP
CC or cell type and to induce antibodies. The dithp nucleic acids are
CC additionally useful in somatic or germline gene therapy of the disorders
CC mentioned above, as a source of antisense sequences, as a source of
CC probes and primers, in genotyping and identification of individuals, in
CC the generation of transgenic animal models of human disease or knock in
CC humanised animals, in toxicological testing, and in transcript imaging.
CC The present sequence represents a DITHP protein which is a cell membrane
CC protein.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 231 AA;

Query Match 38.6%; Score 22; DB 24; Length 231;
Best Local Similarity 36.4%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXVXF 29
| | | | |
Db 105 FTASKNGTYKF 115

RESULT 9
AAG82486 ID AAG82486 standard; Protein; 266 AA.
XX AC AAG82486;
XX DT 03-SEP-2001 (first entry)
XX DE S. epidermidis open reading frame protein sequence SEQ ID NO:2066.
XX
KW Staphylococcus epidermidis SRI strain; infection; diagnosis;
KW vaccination; endocarditis.
XX
OS Staphylococcus epidermidis.
XX
PN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US30782.
XX
PR 09-NOV-1999; 99US-0164258.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Kimmerly WJ;
XX
DR WPI; 2001-316495/33.

DR N-PSDB; AAH53336.
XX
PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis -
XX
PS Claim 18; Page 560; 2188pp; English.
XX
CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG8120, from Staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ Sequence 266 AA;

Query Match 38.6%; Score 22; DB 22; Length 266;
Best Local Similarity 36.4%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXXP 29
| | | | |
Db 147 FKTRSGVYTF 157

RESULT 10
ABP39297
ID ABP39297 standard; Protein; 292 AA.
XX
AC ABP39297;
XX
DT 24-JUL-2002 (first entry)
XX
DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:4142.
XX
KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW antibacterial; gene therapy.
XX
OS Staphylococcus epidermidis.
XX
PN US6380370-B1.
XX
PD 30-APR-2002.
XX
PF 13-AUG-1998; 98US-0134001.
XX
PR 14-AUG-1997; 97US-055779P.
PR 08-NOV-1997; 97US-064964P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Doucette-Stamm LA, Bush D;
XX
DR WPI; 2002-381255/41.
DR N-PSDB; ABN91842.
XX
PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT polypeptide, useful for diagnosing and treating bacterial infections -
XX

PS Disclosure; SEQ ID 4142; 267pp; English.
XX
CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences
CC can also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life
CC cycle or inhibit S. epidermidis infection.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site.
XX
SQ Sequence 292 AA;

Query Match 38.6%; Score 22; DB 23; Length 292;
Best Local Similarity 36.4%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXXP 29
| | | | |
Db 173 FKTRSGVYTF 183

RESULT 11
AAU57054
ID AAU57054 standard; Protein; 540 AA.
XX
AC AAU57054;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #17950.
XX
DE SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59579.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID No 18249; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory

CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 540 AA;

Query Match 38.6%; Score 22; DB 22; Length 540;
Best Local Similarity 36.4%; Pred. No. 6.9e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXXF 29
| | |
Db 160 FSAGSAGDYLF 170

RESULT 12
AAB73384
ID AAB73384 standard; peptide; 20 AA.
XX
AC AAB73384;
XX
DT 05-JUN-2001 (first entry)
XX
DE T. harzianum beta-(1,6)-endoglucanase BGN16.1 N-terminal peptide.
XX
KW Beta-(1,6)-endoglucanase; BGN16.1; strain CECT 2413;
KW antifungal; microbial cell wall degradation; antimicrobial;
KW yeast extract production; disinfection; anti-biofilm;
KW textile treatment; N-terminal peptide.
XX
OS Trichoderma harzianum.
XX
PN WO200109295-A1.
XX
PD 08-FEB-2001.
XX
PF 28-JUL-2000; 2000WO-ES00293.
XX
PR 31-JUL-1999; 99ES-0001747.
XX
PA (NEWB-) NEMBIOTECHNIC SA.
PA (UYSE-) UNIV SEVILLA.
PA (UYSA-) UNIV SALAMANCA.
XX
PI Montero Macarro M, Rey Barrera M, Monte Vazquez E;
PI Llobell Gonzalez A;
XX
DR WPI; 2001-182948/18.
XX
PT New enzyme with beta-(1,6)-endoglucanase activity, useful e.g. as
PT antifungal agent in animals or plants, from Trichoderma harzianum, also
PT related nucleic acid -
XX
PS Example 13; Page 32; 62pp; Spanish.
XX
CC The invention relates to novel Trichoderma harzianum CECT 2413 beta-
CC (1,6)-endoglucanases, designated BGN16.1 and BGN16.3 (AAB73382 and
CC AAB73383, respectively) and to nucleic acids encoding them (AAF76368 and
CC AAF76269, respectively). The enzymes of the invention have antifungal
CC activity, being able to degrade degrade beta-(1,6)-glucan components in
CC microbial cell walls. The beta-(1,6)-endoglucanases are used as
CC antifungal agents for the treatment and/or prevention of infections on
CC plants, animals and humans, and to control species that contaminate foods

CC or harvested crops. They may also be used for disinfection of livestock
CC installations and to prevent contamination of samples for analysis. Other
CC applications of the beta-(1,6)-endoglucanases include the production of
CC protoplasts and yeast extracts; extraction of mannoproteins; production
CC of wine and fruit juices (improving filtration); eliminating dental
CC plaque; cleaning teeth, dentures and contact lenses; removal of biofilms
CC or fungi from coatings; and for the treatment of textiles, e.g., to
CC remove excess dye. The present sequence represents a Trichoderma
CC harzianum CECT 2413 beta-(1,6)-endoglucanase BGN16.1 N-terminal peptide
CC used in the design of a degenerate PCR primer for the isolation of
CC BGN16.1 cDNA.
XX
SQ Sequence 20 AA;

Query Match 36.8%; Score 21; DB 22; Length 20;
Best Local Similarity 36.4%; Pred. No. 49;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXXF 29
| | |
Db 4 FASSQDGRYQF 14

RESULT 13
AAW89961
ID AAW89961 standard; Protein; 111 AA.
XX
AC AAW89961;
XX
DT 20-MAR-2003 (updated)
DT 18-FEB-1999 (first entry)
XX
DE Antigen from cluster 60.
XX
KW Antigen; immunogenic cluster family; vaccine; gastritis; diagnosis;
KW peptic ulcer; gastric adenocarcinoma; gastric lymphoma.
XX
OS Helicobacter pylori.
XX
PN WO9849314-A2.
XX
PD 05-NOV-1998.
XX
PF 25-APR-1998; 98WO-US08487.
XX
PR 14-OCT-1997; 97US-0061958.
PR 25-APR-1997; 97US-0045107.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Chow TP, Fry KE, Lim MY, McAtee CP;
PI WPI; 1999-009433/01.
XX
DR
XX
PT New Helicobacter pylori antigens and related nucleic acid sequences
PT - useful in serological diagnosis and protective vaccines, providing
PT long-lasting immune response
XX
PS Claim 1; Page 272; 402pp; English.
XX
CC The present sequence represents a Helicobacter pylori antigenic protein
CC that is characterised by immunoreactivity with H. pylori-positive
CC antisera. The proteins are highly immunogenic and induce a long-lasting
CC immune response that persists even after antimicrobial treatment. In
CC antibody-detection assays, on sera, plasma, urine, saliva etc., they are
CC highly sensitive and specific. The specification also describes 69
CC previously unrecognised immunogenic cluster families. H. pylori antigens
CC are used to detect H. pylori-specific antibodies, for diagnosing
CC infection or to confirm eradication of infection, and in vaccines to
CC protect against H. pylori infection and related diseases (gastritis,
CC peptic ulcer, gastric adenocarcinoma/lymphoma).
CC (Updated on 20-MAR-2003 to correct PF field.)
XX

SQ Sequence 111 AA;
Query Match 36.8%; Score 21; DB 20; Length 111;
Best Local Similarity 36.4%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 19 FXXXXXGXYYF 29
| | | | |
DB 75 FGAVFTGYAF 85

RESULT 14
ABG76563
ID ABG76563 standard; Protein; 128 AA.
XX
AC ABG76563;
XX
DT 05-NOV-2002 (first entry)
XX
DE HCV E1 antigen monoclonal antibody #51.
XX
KW Human; HCV; hepatitis C virus; E1 antigen; monoclonal antibody; vaccine;
KW hepatotropic; Fab; hypervariable region; E2 antigen; antibody.
XX
OS Homo sapiens.
XX
PN WO200260954-A1.
XX
PD 08-AUG-2002.
XX
PF 14-JAN-2002; 2002WO-SE00044.
XX
PR 12-JAN-2001; 2001US-260889P.
XX
PA (KARO-) KAROLINSKA INNOVATIONS AB.
XX
PI Drakenberg K, Persson MAA;
XX
DR WPI; 2002-608502/65.
XX
PT Vaccine comprising a human monoclonal antibody against hepatitis C
PT virus (HCV) E1 or E2 antigen, useful for treating or preventing HCV
PT infection -
XX
PS Disclosure; Page 55-56; 64pp; English.
XX
CC The invention relates to a human monoclonal antibody or its antigen
CC binding fragments, which exhibit immunological binding affinity for a
CC hepatitis C virus (HCV) E1 antigen, comprising an amino acid sequence
CC homologous to the binding portion of a human antibody Fab molecule from a
CC combinatorial antibody library. The vaccine composition comprising the
CC antibodies or antigen binding fragments against HCV E1 or E2 antigen or
CC its hypervariable region is useful in treating or preventing HCV
CC infection in a subject. Sequences ABG76513-ABG76568 represent human
CC monoclonal antibodies against HCV E1 antigen.
XX
SQ Sequence 128 AA;

Query Match 36.8%; Score 21; DB 23; Length 128;
Best Local Similarity 36.4%; Pred. No. 3.1e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 19 FXXXXXGXYYF 29
| | | | |
DB 105 FVGDTTGYTTF 115

RESULT 15
AAV37071
ID AAV37071 standard; Protein; 151 AA.
XX
AC AAV37071;
XX

DT 07-OCT-1999 (first entry)
XX
DE Protein which is specific to Chlamydia trachomatis.
XX
KW Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;
KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
KW Bartholinitis; pneumopathy; venereal lymphogranulomatosis.
XX
OS Chlamydia trachomatis.
XX
PN WO9928475-A2.
XX
PD 10-JUN-1999.
XX
PF 27-NOV-1998; 98WO-IB01939.
XX
PR 04-NOV-1998; 98US-0107077.
PR 28-NOV-1997; 97FR-0015041.
PR 17-DEC-1997; 97FR-0016034.
XX
PA (GEST) GENSET.
XX
PI Griffais R;
XX
DR WPI; 1999-371125/31.
XX
PT Genome sequence of Chlamydia trachomatis
XX
PS Disclosure; Page 883-884; 1755pp; English.
XX
CC AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
CC of Chlamydia trachomatis (see AA201425). The polypeptides can be used as
CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
CC can also be used to control growth of the microorganism. Chlamydia
CC trachomatis is responsible for a large number of diseases, e.g. eye
CC diseases such as conventional trachoma, nonendemic trachoma,
CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
CC perihepatitis, Bartholinitis; pneumopathy in breast feeding infants;
CC and venereal lymphogranulomatosis. The polypeptides of the invention
CC may be of use in treating these diseases.
XX
SQ Sequence 151 AA;

Query Match 36.8%; Score 21; DB 20; Length 151;
Best Local Similarity 36.4%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 19 FXXXXXGXYYF 29
| | | | |
DB 33 FLSSBSGKYL 43

Search completed: January 12, 2004, 08:15:40
Job time : 6.67166 secs


```
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/030,847
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/039,483
; PRIOR FILING DATE: 1997-02-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-215-088-2
```

```
Query Match      38.6%; Score 22; DB 15; Length 163;
Best Local Similarity 36.4%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```
QY      19 FXXXXXXGXXF 29
      | | | | |
Db      89 FTASKNGTYKF 99
```

RESULT 3

```
; US-09-030-847-4
; Sequence 4, Application US/09030847A
; Patent No. US2002011472A1
; GENERAL INFORMATION:
; APPLICANT: Ni, Jian
; TITLE OF INVENTION: T1-R Ligand III
; FILE REFERENCE: PF357
; CURRENT APPLICATION NUMBER: US/09/030,847A
; CURRENT FILING DATE: 1998-02-26
; EARLIER APPLICATION NUMBER: 60/039,483
; EARLIER FILING DATE: 1997-02-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-030-847-4
```

```
Query Match      38.6%; Score 22; DB 10; Length 215;
Best Local Similarity 36.4%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```
QY      19 FXXXXXXGXXF 29
      | | | | |
Db      89 FTASKNGTYKF 99
```

RESULT 4

```
; US-10-215-088-4
; Sequence 4, Application US/10215088
; Publication No. US20030069409A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: T1-R Ligand III
; FILE REFERENCE: PF357C1
; CURRENT APPLICATION NUMBER: US/10/215,088
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/030,847
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/039,483
; PRIOR FILING DATE: 1997-02-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-215-088-4
```

```
Query Match      38.6%; Score 22; DB 15; Length 215;
Best Local Similarity 36.4%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```
QY      19 FXXXXXXGXXF 29
      | | | | |
Db      89 FTASKNGTYKF 99
```

RESULT 5

```
; US-10-029-386-29069
; Sequence 29069, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29069
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR7.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.75
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.65
; OTHER INFORMATION: SWISSPROT HIT: P34562, EVALU6 6.20e+00
; US-10-029-386-29069
```

```
Query Match      36.8%; Score 21; DB 12; Length 53;
Best Local Similarity 36.4%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```
QY      19 FXXXXXXGXXF 29
      | | | | |
Db      15 FNDNRKGSYSF 25
```

RESULT 6

```
; US-10-039-865-6
; Sequence 6, Application US/10039865
; Publication No. US20020132999A1
; GENERAL INFORMATION:
; APPLICANT: Goli, Surya K.
; APPLICANT: Bandman, Olga
; TITLE OF INVENTION: NOVEL P24 VESICLE PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/039,865
; FILING DATE: 19-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
```

APPLICATION NUMBER: US/09/215,736
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/801,740
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0189 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 201 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1212965
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-039-865-6

Query Match 36.8%; Score 21; DB 14; Length 201;
Best Local Similarity 36.4%; Pred. No. 5.9e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29
DB 82 FAHMDGTYKF 92

RESULT 7
US-10-039-865-7
Sequence 7, Application US/10039865
Publication No. US20020132999A1
GENERAL INFORMATION:
APPLICANT: Goli, Surya K.
BANDMAN, Olga
TITLE OF INVENTION: NOVEL P24 VESICLE PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/039,865
FILING DATE: 19-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/215,736
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/801,740
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0189 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:

LENGTH: 201 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1213221
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-039-865-7

Query Match 36.8%; Score 21; DB 14; Length 201;
Best Local Similarity 36.4%; Pred. No. 5.9e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29
DB 82 FAHMDGTYKF 92

RESULT 8
US-10-060-255-59
Sequence 59, Application US/10060255
Publication No. US20030113840A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 25 Human secreted proteins
FILE REFERENCE: P2042P1
CURRENT APPLICATION NUMBER: US/10/060,255
CURRENT FILING DATE: 2002-02-01
PRIOR FILING DATE: 2001-02-13
PRIOR APPLICATION NUMBER: PCT/US00/22325
PRIOR FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: 60/149,182
PRIOR FILING DATE: 1999-08-17
NUMBER OF SEQ ID NOS: 86
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 59
LENGTH: 201
TYPE: PRT
ORGANISM: Homo sapiens
US-10-060-255-59

Query Match 36.8%; Score 21; DB 15; Length 201;
Best Local Similarity 36.4%; Pred. No. 5.9e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29
DB 82 FAHMDGTYKF 92

RESULT 9
US-10-060-255-68
Sequence 68, Application US/10060255
Publication No. US20030113840A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 25 Human secreted proteins
FILE REFERENCE: P2042P1
CURRENT APPLICATION NUMBER: US/10/060,255
CURRENT FILING DATE: 2002-02-01
PRIOR FILING DATE: 2001-02-13
PRIOR APPLICATION NUMBER: PCT/US00/22325
PRIOR FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: 60/149,182
PRIOR FILING DATE: 1999-08-17
NUMBER OF SEQ ID NOS: 86
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 68
LENGTH: 201
TYPE: PRT

; ORGANISM: Homo sapiens
US-10-060-255-68

Query Match 36.8%; Score 21; DB 15; Length 201;
Best Local Similarity 36.4%; Pred. No. 5.9e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYP 29
| | | | |
DB 82 FAHMDGTYP 92

RESULT 10
US-10-060-255-69

; Sequence 69, Application US/10060255
; Publication No. US20030113840A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 25 Human secreted proteins
; FILE REFERENCE: P2042P1
; CURRENT APPLICATION NUMBER: US/10/060,255
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 09/781,417
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: PCT/US00/22325
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: 60/149,182
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 69
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-255-69

Query Match 36.8%; Score 21; DB 15; Length 201;
Best Local Similarity 36.4%; Pred. No. 5.9e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYP 29
| | | | |
DB 82 FAHMDGTYP 92

RESULT 11
US-09-769-736-86

; Sequence 86, Application US/09769736
; Publication No. US2003013875A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Le Page, Richard WF
; APPLICANT: Wells, Jeremy M
; APPLICANT: Hanniffy, Sean B
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21089WO
; CURRENT APPLICATION NUMBER: US/09/769,736
; PRIOR FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: GB 9816335.5
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/125163
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-09-769-736-86

Query Match 36.8%; Score 21; DB 12; Length 300;
Best Local Similarity 36.4%; Pred. No. 8.6e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYP 29
| | | | |
DB 242 FASVETGTYF 252

RESULT 12
US-10-369-493-11376

; Sequence 11376, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11376
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-11376

Query Match 36.8%; Score 21; DB 12; Length 420;
Best Local Similarity 36.4%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYP 29
| | | | |
DB 201 FALASMGTYF 211

RESULT 13
US-10-369-493-14778

; Sequence 14778, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14778
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-14778

Query Match 36.8%; Score 21; DB 12; Length 420;
Best Local Similarity 36.4%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXYP 29
| | | | |
DB 201 FALASMGTYF 211

RESULT 14

```

US-10-369-493-14932
; Sequence 14932, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14932
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-14932

```

```

Query Match      36.8%; Score 21; DB 12; Length 420;
Best Local Similarity 36.4%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

OY      19 FXXXXXXGXYXF 29
      |  |  |  |  |
DB      201 FALASMGIVVF 211

```

```

RESULT 15
US-10-369-493-15269
; Sequence 15269, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 15269
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-15269

```

```

Query Match      36.8%; Score 21; DB 12; Length 420;
Best Local Similarity 36.4%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

OY      19 FXXXXXXGXYXF 29
      |  |  |  |  |
DB      201 FALASMGIVVF 211

```

Search completed: January 12, 2004, 08:33:42
 Job time : 5.77445 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 08:12:14 ; Search time 1.57784 Seconds
(without alignments)
831.284 Million cell updates/sec

Title: US-10-085-167-3
Perfect score: 57
Sequence: 1 FXXXXXXXXXXXXXXXXXXXXXGXVFX 31

Scoring table: Biosum62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep: *
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep: *
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep: *
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep: *
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep: *
6: /cgn2_6/ptodata/2/iaa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------------------|-------------------|
| 1 | 22 | 38.6 | 268 | 4 | US-09-252-991A-18273 | Sequence 18273, A |
| 2 | 22 | 38.6 | 292 | 4 | US-09-134-001C-4142 | Sequence 4142, Ap |
| 3 | 22 | 38.6 | 305 | 4 | US-09-328-352-6069 | Sequence 6069, Ap |
| 4 | 22 | 38.6 | 404 | 4 | US-09-328-352-8182 | Sequence 8182, Ap |
| 5 | 21 | 36.8 | 156 | 4 | US-09-198-452A-698 | Sequence 698, App |
| 6 | 21 | 36.8 | 201 | 2 | US-08-801-740-6 | Sequence 6, Appli |
| 7 | 21 | 36.8 | 201 | 2 | US-08-801-740-7 | Sequence 7, Appli |
| 8 | 21 | 36.8 | 201 | 3 | US-08-801-740-6 | Sequence 6, Appli |
| 9 | 21 | 36.8 | 201 | 3 | US-08-801-740-7 | Sequence 7, Appli |
| 10 | 21 | 36.8 | 252 | 4 | US-09-252-991A-30813 | Sequence 30813, A |
| 11 | 21 | 36.8 | 1416 | 4 | US-09-071-035-404 | Sequence 404, App |
| 12 | 21 | 36.8 | 1448 | 4 | US-09-071-035-402 | Sequence 402, App |
| 13 | 20 | 35.1 | 63 | 2 | US-08-637-759B-501 | Sequence 501, App |
| 14 | 20 | 35.1 | 63 | 3 | US-08-871-355A-501 | Sequence 501, App |
| 15 | 20 | 35.1 | 63 | 4 | US-09-201-945-501 | Sequence 501, App |
| 16 | 20 | 35.1 | 142 | 2 | US-08-997-080-47 | Sequence 47, Appl |
| 17 | 20 | 35.1 | 142 | 2 | US-08-997-362-47 | Sequence 47, Appl |
| 18 | 20 | 35.1 | 142 | 3 | US-08-873-970-47 | Sequence 47, Appl |
| 19 | 20 | 35.1 | 142 | 3 | US-09-095-855-47 | Sequence 47, Appl |
| 20 | 20 | 35.1 | 142 | 3 | US-08-705-347A-47 | Sequence 47, Appl |
| 21 | 20 | 35.1 | 142 | 4 | US-09-324-542-47 | Sequence 47, Appl |
| 22 | 20 | 35.1 | 142 | 4 | US-09-205-426-47 | Sequence 47, Appl |
| 23 | 20 | 35.1 | 142 | 4 | US-09-200-643-47 | Sequence 47, Appl |
| 24 | 20 | 35.1 | 143 | 3 | US-08-685-808-2 | Sequence 2, Appli |
| 25 | 20 | 35.1 | 143 | 3 | US-08-505-860C-2 | Sequence 2, Appli |
| 26 | 20 | 35.1 | 162 | 2 | US-08-606-143-44 | Sequence 44, Appl |
| 27 | 20 | 35.1 | 336 | 4 | US-08-987-691A-5 | Sequence 5, Appli |

| | | | | | | |
|----|----|------|------|---|----------------------|-------------------|
| 28 | 20 | 35.1 | 408 | 4 | US-09-252-991A-33131 | Sequence 33131, A |
| 29 | 20 | 35.1 | 469 | 4 | US-09-198-452A-793 | Sequence 793, App |
| 30 | 20 | 35.1 | 479 | 4 | US-09-996-243-216 | Sequence 216, App |
| 31 | 20 | 35.1 | 492 | 4 | US-09-252-991A-27853 | Sequence 27853, A |
| 32 | 20 | 35.1 | 498 | 3 | US-09-232-468A-18 | Sequence 18, Appl |
| 33 | 20 | 35.1 | 498 | 3 | US-09-232-468A-24 | Sequence 24, Appl |
| 34 | 20 | 35.1 | 498 | 3 | US-08-686-968C-231 | Sequence 231, App |
| 35 | 20 | 35.1 | 498 | 4 | US-09-784-984B-52 | Sequence 52, Appl |
| 36 | 20 | 35.1 | 498 | 4 | US-09-784-984B-54 | Sequence 54, Appl |
| 37 | 20 | 35.1 | 546 | 2 | US-09-067-351-1 | Sequence 1, Appli |
| 38 | 20 | 35.1 | 546 | 3 | US-09-360-490-1 | Sequence 1, Appli |
| 39 | 20 | 35.1 | 646 | 4 | US-09-336-643A-10 | Sequence 10, Appl |
| 40 | 20 | 35.1 | 922 | 4 | US-09-198-452A-15 | Sequence 15, Appl |
| 41 | 20 | 35.1 | 1430 | 3 | US-09-008-172-2 | Sequence 2, Appli |
| 42 | 20 | 35.1 | 1430 | 3 | US-09-210-361-6 | Sequence 6, Appli |
| 43 | 20 | 35.1 | 1430 | 4 | US-09-740-274-6 | Sequence 6, Appli |
| 44 | 20 | 35.1 | 3177 | 2 | US-08-477-451-4 | Sequence 4, Appli |
| 45 | 19 | 33.3 | 32 | 4 | US-08-525-539A-35 | Sequence 35, Appl |

ALIGNMENTS

```
RESULT 1
US-09-252-991A-18273
; Sequence 18273, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252, 991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074, 788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094, 190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18273
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18273

Query Match      38.6%; Score 22; DB 4; Length 268;
Best Local Similarity 36.4%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      19 FXXXXXXXXXXF 29
Db      41 FLAVATGYAF 51

RESULT 2
US-09-134-001C-4142
; Sequence 4142, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134, 001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064, 964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055, 779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4142
; LENGTH: 292
; TYPE: PRT
```

; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4142

Query Match 38.6%; Score 22; DB 4; Length 292;
Best Local Similarity 36.4%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXXF 29
| | | | |
Db 173 FKTRSGVTYF 183

RESULT 3
US-09-328-352-6069
; Sequence 6069, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6069
; LENGTH: 305
; TYPE: PRT

; ORGANISM: Acinetobacter baumannii
US-09-328-352-6069

Query Match 38.6%; Score 22; DB 4; Length 305;
Best Local Similarity 36.4%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXXF 29
| | | | |
Db 52 FTGTAQGTYYF 62

RESULT 4
US-09-328-352-8182

; Sequence 8182, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8182
; LENGTH: 404
; TYPE: PRT

; ORGANISM: Acinetobacter baumannii
US-09-328-352-8182

Query Match 38.6%; Score 22; DB 4; Length 404;
Best Local Similarity 36.4%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXXF 29
| | | | |
Db 242 FLKTTAGHYAF 252

RESULT 5
US-09-198-452A-698

; Sequence 698, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments

; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention

; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 698
; LENGTH: 156
; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-698

Query Match 36.8%; Score 21; DB 4; Length 156;
Best Local Similarity 36.4%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXXF 29
| | | | |
Db 88 FSLPSSGTYVF 98

RESULT 6
US-08-801-740-6

; Sequence 6, Application US/08801740
; Patent No. 5869639
; GENERAL INFORMATION:

; APPLICANT: Goli, Surya K.
; APPLICANT: Bandman, Olga
; TITLE OF INVENTION: NOVEL P24 VESICLE PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,740
; FILING DATE: Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:

; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0189 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:

; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1212965
; US-08-801-740-6

Query Match 36.8%; Score 21; DB 2; Length 201;
Best Local Similarity 36.4%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXGXXF 29

Db 82 FAAHMDGTYKF 92

RESULT 7

US-08-801-740-7
; Sequence 7, Application US/08801740
; Patent No. 5869639

GENERAL INFORMATION:

APPLICANT: Goli, Surya K.
APPLICANT: Bandman, Olga
TITLE OF INVENTION: NOVEL P24 VESICLE PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,740
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0189 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:
LENGTH: 201 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1213221
US-08-801-740-7

Query Match 36.8%; Score 21; DB 2; Length 201;
Best Local Similarity 36.4%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXXGXXF 29
Db 82 FAAHMDGTYKF 92

RESULT 8

US-08-801-740-6
; Sequence 6, Application US/08801740
; Patent No. 6130325

GENERAL INFORMATION:

APPLICANT: Goli, Surya K.
APPLICANT: Bandman, Olga
TITLE OF INVENTION: NOVEL P24 VESICLE PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto

STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,740
FILING DATE: 14-Feb-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0189 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 201 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1212965
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-08-801-740-6

Query Match 36.8%; Score 21; DB 3; Length 201;
Best Local Similarity 36.4%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXXGXXF 29
Db 82 FAAHMDGTYKF 92

RESULT 9

US-08-801-740-7
; Sequence 7, Application US/08801740
; Patent No. 6130325

GENERAL INFORMATION:

APPLICANT: Goli, Surya K.
APPLICANT: Bandman, Olga
TITLE OF INVENTION: NOVEL P24 VESICLE PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,740
FILING DATE: 14-Feb-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0189 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 201 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1213221
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-08-801-740-7

Query Match 36.8%; Score 21; DB 3; Length 201;
Best Local Similarity 36.4%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXXGXYXF 29
DB 82 FAAHMDGTYKF 92

RESULT 10
US-09-252-991A-30813
Sequence 30813, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 30813
LENGTH: 252
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30813

Query Match 36.8%; Score 21; DB 4; Length 252;
Best Local Similarity 36.4%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXXGXYXF 29
DB 167 FASEDSGVYRF 177

RESULT 11
US-09-071-035-404
Sequence 404, Application US/09071035
Patent No. 6448043
GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA

ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 404:
SEQUENCE CHARACTERISTICS:
LENGTH: 1416 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-404

Query Match 36.8%; Score 21; DB 4; Length 1416;
Best Local Similarity 36.4%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 19 FXXXXXXGXYXF 29
DB 959 FTTDEKQYSF 969

RESULT 12
US-09-071-035-402
Sequence 402, Application US/09071035
Patent No. 6448043
GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 402:

SEQUENCE CHARACTERISTICS:
LENGTH: 1448 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-402

Query Match 36.8%; Score 21; DB 4; Length 1448;
Best Local Similarity 36.4%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29
DB 986 FTTDEKGYYSF 996

RESULT 13
US-08-637-759B-501
Sequence 501, Application US/08637759B
Patent No. 5876931
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/637,759B
FILING DATE: 03-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02875
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 501:
SEQUENCE CHARACTERISTICS:
LENGTH: 63 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-637-759B-501

Query Match 35.1%; Score 20; DB 2; Length 63;
Best Local Similarity 36.4%; Pred. No. 85;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29
DB 4 FQSIYAGYYF 14

RESULT 14

US-08-871-355A-501
Sequence 501, Application US/08871355A
Patent No. 6015669
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/871,355A
FILING DATE: 09-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02875
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101 CON
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 501:
SEQUENCE CHARACTERISTICS:
LENGTH: 63 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-871-355A-501

Query Match 35.1%; Score 20; DB 3; Length 63;
Best Local Similarity 36.4%; Pred. No. 85;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 19 FXXXXXXGXXF 29
DB 4 FQSIYAGYYF 14

RESULT 15
US-09-201-945-501
Sequence 501, Application US/09201945
Patent No. 6342215
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/201,945
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/637,759
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPLS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 501:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 63 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-09-201-945-501

```

```

Query Match      35.1%; Score 20; DB 4; Length 63;
Best Local Similarity 36.4%; Pred. No. 85;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

QY      19 FXXXXXGXFXF 29
      | | | | |
DB      4 FQSIIVAGYYF 14

```

Search completed: January 12, 2004, 08:20:09
 Job time : 3.57784 secs